Elektronica en Telecommunicatie 9

Gemeenschappelijke publikatie van de
Sectie voor Telecommunicatietechniek van het K.I.v.I. en het
Nederlands Elektronica- en Radiogenootschap.
Redactie-adres: Prinsessegracht 23, Den Haag.

681.327.12

Patroonherkenning

I. Het patroonherkennend systeem en enkele hierin toepasbare technieken

door ir. R. P. W. Duin, Laboratorium voor Technische Natuurkunde, T.H. Delft



Synopsis: *The Pattern Recognition System and some General Applicable Techniques.*
The pattern recognition system is considered as a system that can classify into known classes those objects, which are physically observable phenomena. Only the case is discussed in which the classes can be defined by objects of known classification.

A model for classifying systems is given. This model consists of three parts: a *transducer*, which performs measurements on the objects, a *pre-processor*, which selects features from the measurements with the help of 'a priori' knowledge and heuristic algorithms, and a *decision unit*, which classifies the objects according to the features.

Some techniques for the construction of pre-processors and decision units are discussed. Finally some applications and a brief prospect are given.

1. Inleiding

Er bestaat in het gebied van patroonherkennen een grote variëteit aan problemen, technieken en toepassingen. We zullen ons in dit artikel tot enkele algemene aspecten beperken. Enkele deelgebieden, die soms wel tot patroonherkennen worden gerekend, zullen buiten beschouwing blijven. Hiertoe behoort o.a. het herkenningsmechanisme in biologische systemen. Alleen kunstmatige systemen zullen worden bekeken. Door Nagy (1968) is een overzicht van vele problemen en technieken gegeven [9].

Het ingangssignaal van een patroonherkennend systeem wordt in het algemeen gevormd door een aantal metingen aan objecten. Objecten zijn fysisch waarneembare grootheden (bijv. 'signalen' als gesproken woord en elektrocardiogrammen, of 'beelden' als cijfers en auto's). Het uitgangssignaal bestaat uit een classificatie van het aangeboden object. Deze classificatie is een uitspraak omtrent het al dan niet behoren tot een bepaalde klasse van objecten. Het is mogelijk om m.b.v. dit systeem andere systemen te regelen. Er kan dan sprake zijn van zogenaamd *lerend regelen* (learning control). Dergelijke aspecten van patroonherkennen zullen niet in beschouwing worden genomen. Overzichtsartikelen op dit gebied zijn gepubliceerd door Sklansky (1966) [12] en Fu (1970) [4].

De klassen waarin objecten worden verdeeld kunnen op drie manieren worden gedefinieerd:

a. *Door het systeem zelf.* Op grond van de aangeboden objecten gaat het systeem klassen onderscheiden. Het is mogelijk dat het voortdurend blijft leren, dus dat het, wanneer daar aanleiding toe is, de klassenindeling wijzigt. Vaak wordt het echter in dezelfde toestand gehouden als het na een zekere leerperiode bevredigend werkt. Aangetekend moet nog worden dat het bij het ontwerpen van het systeem vaak noodzakelijk is om voorkennis omtrent de objecten te gebruiken.

We zullen de systemen van deze groep verder buiten beschouwing laten. Ze zijn o.a. beschreven door Patrick (1968) [11] en Cooper (1969) [2].

Voordrachten, gehouden tijdens de 214e werkvergadering van het NERG op 29 januari 1971, in het gebouw van het Koninklijk Instituut van Ingenieurs te 's-Gravenhage.

b. *Met behulp van leerobjecten*. Van elke der te onderscheiden klassen wordt een aantal objecten met bekende classificatie (*gelabelde objecten*) gebruikt om het systeem te ontwerpen en zodanig in te stellen dat de klassen zo goed mogelijk worden onderscheiden. Na deze leerfase kan het systeem operationeel worden. Er vinden dan geen wijzigingen meer in plaats.

c. *Met behulp van vooraf bepaalde, meetbare criteria*. In dit geval is precies bekend waaraan de metingen aan de objecten van de diverse klassen behoren te voldoen. We zullen dergelijke gevallen, die zich slecht lenen voor algemene beschouwingen, verder niet in aanmerking nemen.

We beperken ons verder tot het onder b. genoemde geval, waarbij de klassen dus impliciet door een aantal gelabelde objecten worden gedefinieerd.

2. Het patroonherkennend systeem

Het totale classificatieproces kan worden beschreven als een vorm van informatiereductie. Uitgaande van alle gegevens die van een object zijn te bepalen blijft uiteindelijk alleen de classificatie over. Er zijn drie methoden om deze informatiereductie te bewerkstelligen:

a. *Met behulp van a priori kennis*. Door gebruik te maken van kennis die men van objecten en klassen heeft, kunnen metingen en combinaties van metingen worden geselecteerd die vermoedelijk relevant zijn voor de classificatie.

b. *Met behulp van een statistische analyse van de leerobjecten.*

c. *Met heuristische methoden*. Hierbij wordt vaak weer gebruik gemaakt van statistische gegevens en a priori kennis.

De resultaten van de laatste twee methoden zijn in het algemeen beter wanneer het aantal parameters kleiner is. Het verdient daarom aanbeveling om in eerste instantie zoveel mogelijk van de a priori kennis gebruik te maken.

Op grond hiervan is het in fig. 1 geschetste model voor het patroonherkennend systeem opgesteld. De 'opnemer' meet *die* gegevens aan het object waarvan vermoed wordt, dat ze relevant zijn. Op grond van a priori kennis, en soms met behulp van heuristische methoden wordt een 'bewerker' geconstrueerd die zoveel mogelijk niet-relevante informatie elimineert. Het uitgangssignaal van de bewerker bestaat uit zgn. kenmerken. Dit zijn combinaties en transformaties van metingen, waarin de relevante informatie zoveel mogelijk is geconcentreerd. Tot slot wordt m.b.v. kenmerken de beslissing genomen omtrent de classificatie.

Een voordeel van een dergelijk model is dat bij de constructie van de 'beslisser' geen gebruik gemaakt behoeft te worden van a priori kennis over de objecten waarvoor het systeem wordt opgezet. Hierdoor is het mogelijk beslissers in het algemeen te bestuderen. In zo'n algemene beslisser komen vaak nog een aantal parameters voor. Deze kunnen m.b.v. de leerobjecten worden ingesteld.

Wanneer het systeem eenmaal is geconstrueerd kan het worden getest met een aantal gelabelde objecten. Dit dienen uiteraard andere dan de leerobjecten te zijn. Wanneer de werking niet bevredigend is, kan worden geprobeerd meer leerobjecten te nemen. Hierdoor kan een beter gebruik gemaakt worden van statistische gegevens. Een andere mogelijkheid is een nadere studie van de objecten en de klassen te maken. De a priori kennis wordt hierdoor vergroot, zodat wellicht een betere bewerker kan worden gemaakt. Het gegeven model met een bewerker en een beslisser zal niet voor alle patroonherkennende systemen voldoen. Soms is de beslisser praktisch verdwenen en wordt alles met behulp van de a priori kennis opgelost. Dit is onder meer het geval bij systemen die een duidelijke simulatie van het menselijk beslissingsproces in een specifiek geval vormen.

De functies en de mogelijkheden van de bewerker en de beslisser zullen in het nu volgende wat verder worden uitgewerkt.

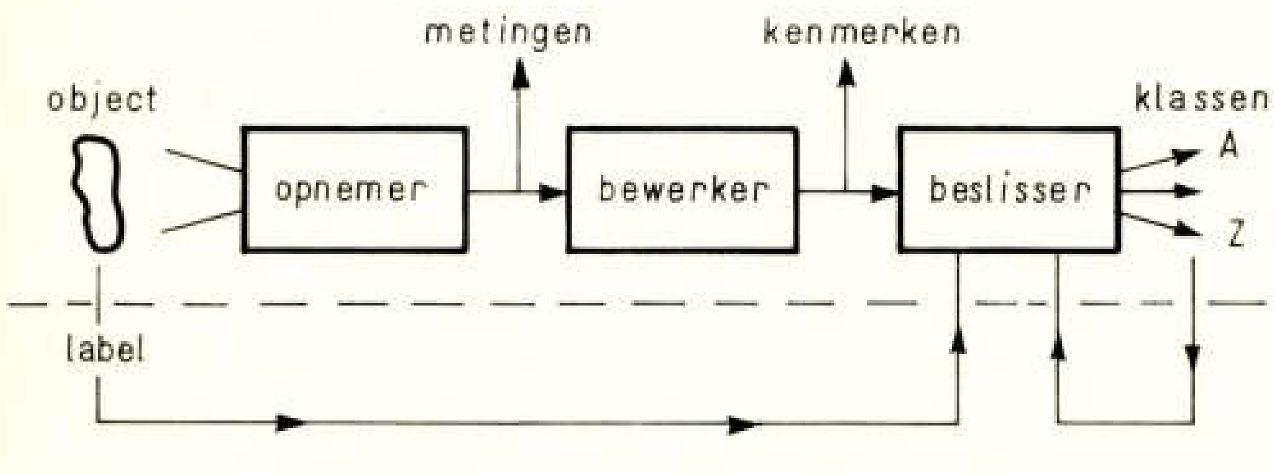


Fig. 1. Een model voor het patroonherkennend systeem.

3. De bewerker

De bewerker heeft twee taken. Allereerst dient hij zo veel mogelijk van de niet-relevante informatie in het opnemersignaal te elimineren. Bovendien moet de informatie in de vorm van voor de beslisser geschikte kenmerken worden afgegeven. Gezien het voorgaande dient het aantal kenmerken zo klein mogelijk te zijn. De kenmerken kunnen van kwantitatieve aard zijn, zoals bijvoorbeeld continu variabele parameters, of van kwalitatieve aard, zoals het al dan niet aanwezig zijn van een bepaalde eigenschap. Een en ander legt uiteraard eisen op aan de beslisser.

Algemene methoden voor het bepalen van kenmerken uit een aantal meetgegevens zijn er vrijwel nog niet. Voor bepaalde typen objecten, zoals eenvoudige tweedimensionale figuren, bestaan er wel een aantal standaardtechnieken, zie Levine (1969) [8]. We zullen nu enige mogelijkheden noemen, volgens welke relevante kenmerken kunnen worden opgespoord.

a. *Op heuristische gronden kunnen metingen worden geselecteerd, gecombineerd en getransformeerd*. Hierbij wordt gebruik gemaakt van de beschikbare a priori kennis van de objecten. Voorbeelden zijn het zoeken van eindpunten, drie- en viersprongen e.d. bij het herkennen van schrift, of de bepaling van hoogte en lengte bij het onderscheiden van vrachtauto's en personenauto's. Wanneer er aanleiding toe is kunnen ook fouriertransformaties en polynoombenaderingen van signalen worden geprobeerd.

b. *Het genereren en proberen van combinaties van metingen*. In sommige gevallen, bijv. wanneer het aantal mogelijke combinaties niet te groot is, kan het zin hebben om alle, of een aantal combinaties te proberen. Ook is het mogelijk om volgens een op heuristische gronden geconstrueerd patroon combinaties te genereren en te proberen. Dit is o.a. beschreven door Bakis, Herbst en Nagy (1968) [1]. Het nagaan van de aanwezigheid van correlaties tussen metingen is van belang bij het zoeken naar geschikte combinaties.

c. *Clusteranalyse.* Is eenmaal een aantal kenmerken geselecteerd, dan kan worden nagegaan in hoeverre er op grond van deze kenmerken structuur binnen een klasse aanwezig is. Wanneer namelijk blijkt dat een klasse duidelijk in een aantal subklassen uiteenvalt, kan het voordelen hebben elk van de subklassen afzonderlijk te beschouwen in plaats van de totale klasse. Er zijn echter nog geen methoden bekend die met grote hoeveelheden data in alle opzichten bevredigend werken.

d. *Dimensieanalyse.* Hieronder wordt verstaan het onderzoek naar het aantal dimensies dat wezenlijk van belang is bij de beschrijving van een groep objecten. Wanneer alle objecten in een, eventueel niet-lineaire, deelruimte van de kenmerkruimte (zie par. **4**) zouden blijken te liggen, heeft het zin te zoeken naar de kenmerken die deze deelruimte beschrijven. Hierdoor wordt een kenmerkreductie bereikt.

Een probleem bij het bepalen van geschikte kenmerken is dat het meestal niet mogelijk is een kenmerk afzonderlijk op zijn waarde te toetsen. Dit komt doordat de kenmerken in de beslisser samenwerken. Hierdoor is het mogelijk dat kenmerken samen betere resultaten leveren dan op grond van de resultaten van de kenmerken afzonderlijk zou worden verwacht. Zie bijv. fig. 2, waar vanwege de correlatie tussen de kenmerken een betere scheiding mogelijk is dan uit de kenmerken afzonderlijk zou blijken. Uitsluitend in het geval, waarin de kenmerken ongecorreleerd zijn, heeft het zin ze afzonderlijk op hun bruikbaarheid te toetsen.

Een criterium hiervoor, wanneer scheiding van twee klassen wordt beoogd, is het criterium van Fisher:

$$L = \frac{(\mu_A - \mu_B)^2}{\sigma_A^2 + \sigma_B^2}$$

waarin voor het onderzochte kenmerk μ_A en μ_B de gemiddelde waarden en σ_A^2 en σ_B^2 de varianties van de beide klassen zijn.

Een grote waarde van L duidt op een grote scheidbaarheid. Dit is echter alleen zinvol als de beslisser een lineaire scheidingsfunctie (zie par. **4**) instelt. In kwadratisch scheidbare situaties kan de waarde van L toch nul worden, hoewel dit zou duiden op zeer slechte scheidbaarheid.

Wanneer men de beschikking heeft over een vast aantal leerobjecten waaraan in principe zeer veel verschillende metingen gedaan kunnen worden, is het gevaarlijk om al deze metingen

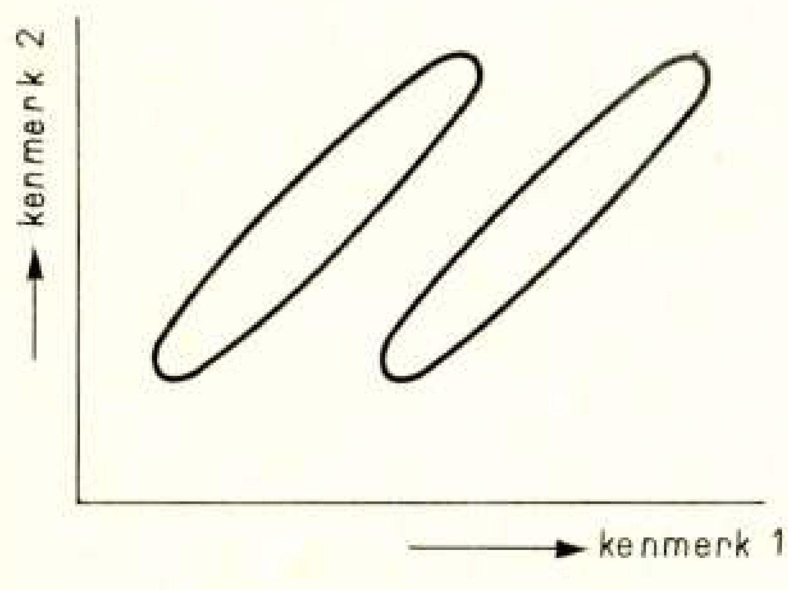


Fig. 2. Voorbeeld van de ligging van twee klassen waarbij de kenmerken afzonderlijk verkeerd zouden kunnen worden gewaardeerd ten gevolge van de correlatie tussen de kenmerken. Wanneer in dit geval alleen op kenmerk 2 zou worden gelet zou men ten onrechte kunnen menen dat de bijdrage van dit kenmerk nihil is.

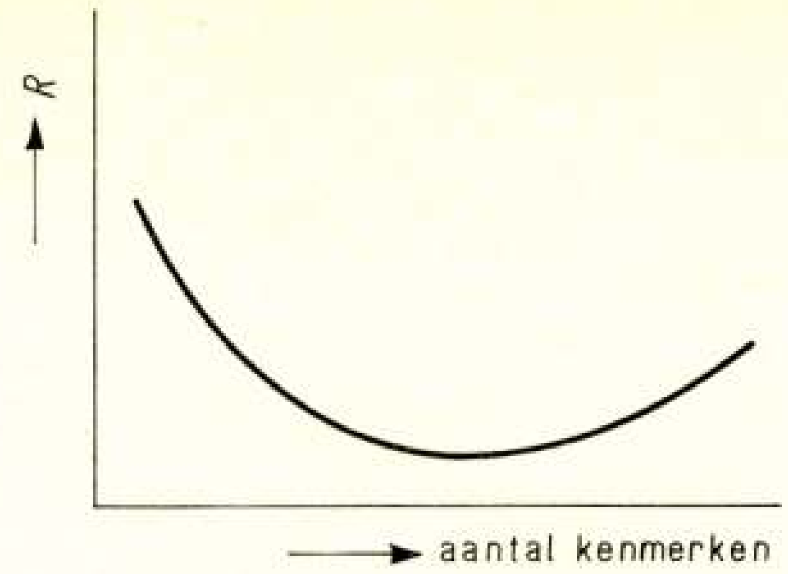


Fig. 3. Verloop van de kans op foutclassificaties R als functie van het aantal kenmerken bij een constant aantal leerobjecten.

maar als kenmerken te gebruiken. In fig. 3 is het verloop geschetst van de fouten, die met een gevonden scheidingsfunctie uiteindelijk kunnen worden gemaakt, als functie van het gebruikte aantal kenmerken, en bij een vast aantal leerobjecten. Bij toeneming van het aantal kenmerken zullen de resultaten eerst verbeteren, maar uiteindelijk verslechteren. Dit geldt overigens alleen wanneer de nieuw toegevoegde kenmerken niet beter zijn dan de reeds bestaande. De oorzaak hiervan is waarschijnlijk dat het aantal gegevens op grond waarvan de beslisser wordt ingesteld, lineair toeneemt met het aantal kenmerken. Deze gegevens zijn de kengetallen van alle leerobjecten. De grootte van de kenmerkruimte, waarin de beslissing uiteindelijk moet worden genomen, stijgt echter met de macht van het aantal kenmerken.

4. De beslisser

Bij de behandeling van de beslisser zullen wij ons twee beperkingen opleggen. Wij zullen alleen het geval behandelen waarin onderscheid wordt gemaakt tussen slechts twee klassen, en bovendien aannemen dat de kenmerken continu variabel zijn. Wanneer ze discreet of binair zijn zal een aantal genoemde methoden toch, zij het na enige modificatie, bruikbaar zijn. In het geval dat de kenmerken kwalitatief van aard zijn wordt evenwel een geheel andere aanpak dan de behandelde vereist.

We zullen de ingang van de beslisser, de verzameling van kenmerken, voorstellen door een vector $\underline{x}$. Door de kenmerken wordt een zgn. kenmerkruimte opgespannen. De leerobjecten van de klassen A en B worden voorgesteld door vectoren in deze ruimte, $\underline{x}_{A_i}$ ($i = 1,2,\ldots,m$) resp. $\underline{x}_{B_i}$ ($i = 1,2,\ldots,n$). De beslissing omtrent de classificatie van een object $\underline{x}$ wordt genomen m.b.v. een scheidingsfunctie $F(\underline{x})$, en wel volgens (zie fig. 4):

$F(\underline{x}_o) > 0$ $\quad \underline{x}_o$ uit klasse A

$F(\underline{x}_o) = 0$ $\quad \underline{x}_o$ uit klasse A of B

$F(\underline{x}_o) < 0$ $\quad \underline{x}_o$ uit klasse B

Wij zullen twee typen methoden onderscheiden om $F(\underline{x})$ te vinden, nl. de statistische methoden en de heuristische methoden.

4.1. *Statistische methoden*

Deze methoden berusten op de strategie dat $F(\underline{x})$ zodanig moet worden bepaald dat, wanneer ermee wordt geclassificeerd, de kosten veroorzaakt door optredende foutclassificaties minimaal zijn. Wanneer voor de objecten uit klasse A een kansdichtheids-

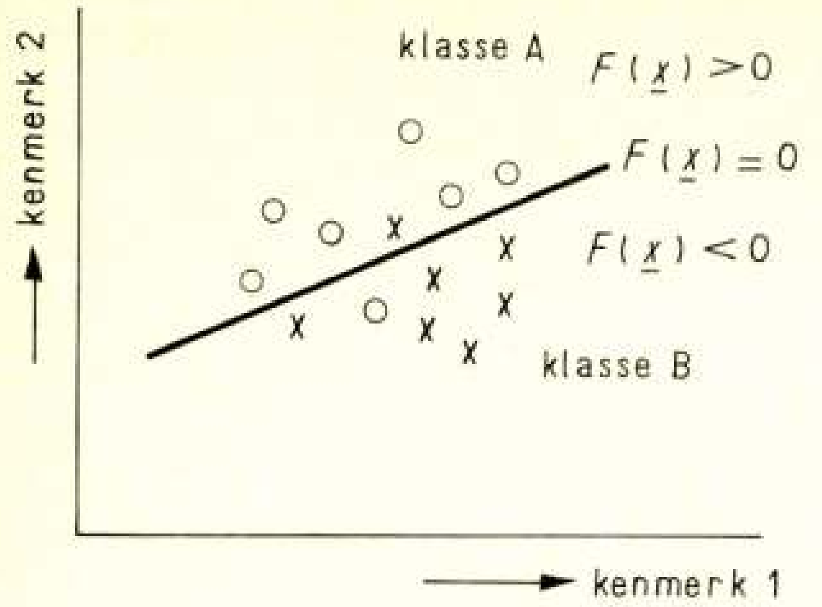


Fig. 4. Voorbeeld van een scheidingsfunctie in de kenmerkruimte.

verdeling $P_A(\underline{x})$ en voor de objecten uit klasse B een kansdichtheidsverdeling $P_B(\underline{x})$ geldt, dan zijn de kosten die bij gebruik van $F(\underline{x})$ worden gemaakt:

$$h_A l_A \int\limits_{F(\underline{x})<0} P_A(\underline{x})\,\mathrm{d}\underline{x} + h_B l_B \int\limits_{F(\underline{x})>0} P_B(\underline{x})\,\mathrm{d}\underline{x} \qquad (1)$$

Hierin zijn h_A en h_B de a priori waarschijnlijkheden van het voorkomen van objecten uit resp. klasse A en B, en l_A en l_B de kosten die verbonden zijn aan een foutclassificatie van een object van klasse A resp. B. Vrij eenvoudig is te bewijzen dat (1) minimaal is als:

$$F(\underline{x}) = h_A l_A P_A(\underline{x}) - h_B l_B P_B(\underline{x}) \qquad (2)$$

De statistische methoden zijn gebaseerd op het schatten van de kansdichtheidsverdelingen m.b.v. de leerobjecten, waarna uit (2) $F(\underline{x})$ volgt. Het schatten van $P_A(\underline{x})$ en $P_B(\underline{x})$ kan op twee manieren gebeuren:

a. *Parametrisch.* Aangenomen wordt dat de verdelingen van een bekend type zijn. Met behulp van de leerobjecten worden nu de parameters van de verdelingen geschat. In het geval van een normale verdeling zijn dit bijvoorbeeld de covariantiematrix en het gemiddelde.

b. *Non-parametrisch.* Een andere manier om de verdelingen $P_A(\underline{x})$ en $P_B(\underline{x})$ uit de leerobjecten te schatten is ieder leerobject vervangen te denken door een potentiaalfunctie $\phi(\underline{x})$ en vervolgens over de leerobjecten te sommeren:

$$\hat{P}_A(\underline{x}) = \frac{1}{m} \sum_{i=1}^{m} \phi(\underline{x} - \underline{x}_{A_i})$$

$$\hat{P}_B(\underline{x}) = \frac{1}{n} \sum_{i=1}^{n} \phi(\underline{x} - \underline{x}_{B_i})$$

Bekende keuzen voor ϕ zijn:

$$\exp\left(-\frac{\underline{x}\cdot\underline{x}}{2\sigma^2}\right) \text{ en } \frac{1}{1+\underline{x}\cdot\underline{x}}$$

Door Specht (1966) [13] is een methode ontwikkeld, die op de eerstgenoemde functie is gebaseerd.

4.2. *Heuristische methoden*

Tussen de diverse methoden uit deze groep bestaat weinig samenhang.

4.2.1. *Minimalisering van het aantal foutgeclassificeerde leerobjecten.*

Hierbij wordt voor $F(\underline{x})$ een bepaalde vorm gekozen, bijv. $F(\underline{x}) = \underline{w} \cdot \underline{x} + w_o$. De parameters $\underline{w}$ en w_o dienen nu zo te worden bepaald, dat het classificatieresultaat van de leerobjecten met $F(\underline{x})$ zo goed mogelijk is. Het optimaliseren van $F(\underline{x})$ kan gebeuren door alle leerobjecten tegelijk in aanmerking te nemen (simultane methoden), of door de classificaties van de leerobjecten stuk voor stuk te bekijken en aan de hand hiervan eventueel correcties aan $F(\underline{x})$ aan te brengen (sequentiële methoden).

a. *Simultane methoden.* Voor $F(\underline{x})$ moet nu gelden dat de leerobjecten zo goed mogelijk worden geclassificeerd, dus dat:

$$\frac{1}{m} \sum_{i=1}^{m} \mathrm{wgf}\,(F(\underline{x}_{A_i})) + \frac{1}{n} \sum_{i=1}^{n} \mathrm{wgf}\,(-F(\underline{x}_{B_i}))$$

minimaal is. Hierin is $\mathrm{wgf}(F(\underline{x}_{A_i}))$ een weegfunctie van het classificatieresultaat van het leerobject $\underline{x}_{A_i}$. Wanneer $F(\underline{x})$ een genormeerde functie is wordt in feite de afstand van $\underline{x}_{A_i}$ tot het scheidingsvlak $F(\underline{x}) = 0$ gewogen. De meest voor de hand liggende weegfunctie is de stapfunctie:

$$\mathrm{wgf}(y) = 0 \text{ voor } y \geqq 0$$

$$\mathrm{wgf}(y) = 1 \text{ voor } y < 0$$

Hierdoor wordt het aantal foutgeclassificeerde leerobjecten geminimaliseerd. Om het minimaliseringsproces wat makkelijker te laten verlopen worden vaak continu verlopende functies gebruikt. Problemen die optreden zijn de keuze van de weegfunctie wgf en het minimaliseringsproces. Aangetoond kan worden dat voor sommige weegfuncties deze methode dezelfde is als de eerder genoemde non-parametrische statistische methode met $h_A l_A = h_B l_B$.

b. *Sequentiële methoden.* Met de classificatie van elk leerobject afzonderlijk wordt de scheidingsfunctie $F(\underline{x})$ via $\underline{w}$ en w_o bijgewerkt. De correctie is in het algemeen een functie van het classificatieresultaat, van de gewenste classificatie van het betreffende leerobject $\underline{x}_i$, en van $F(\underline{x}_i)$. Er bestaat een grote variëteit aan methoden op dit gebied, waarvan sommige convergentie garanderen en andere niet. Het belang van deze methoden is dat ze relatief eenvoudig zijn te instrumenteren, daar ze geen snel toegankelijke geheugens voor de leerobjecten vereisen. *Adaline* (*ada*ptive *lin*ear *e*lement) en *TLU* (*t*hreshold *l*inear *u*nit) zijn benamingen voor apparaten die volgens een sequentiële methode werken, zie Nilsson (1965) [10].

4.2.2. *Methode van de convex omhullende*

Van de leerobjecten der beide klassen A en B wordt de convex omhullende bepaald. Hiermee wordt de ruimte in vier gebieden verdeeld, nl. een gebied dat tot klasse A hoort, een gebied dat tot klasse B hoort, een gebied dat tot beide klassen hoort en een gebied dat tot geen van beide klassen hoort. Hiermee kunnen objecten worden ingedeeld. Vallen ze buiten de beide klassen, of in het overlappingsgebied, dan kan een willekeurige keuze worden gedaan, of niet-automatisch worden ingedeeld.

4.2.3. *De naaste-buurmethode*

Hierbij wordt een object toegewezen aan de klasse van het leerobject dat de kleinste afstand tot het aangeboden object heeft. De praktische bruikbaarheid van deze methode is gering, daar het noodzakelijk is om voor elk aangeboden object de afstanden tot alle leerobjecten te berekenen. Het belang van deze methode is echter dat zij een indruk geeft over de best mogelijke scheiding, zie bijv. Cover en Hart (1967) [3].

4.2.4. *De sjabloonmethode*

Bij deze methode wordt van een aangeboden signaal de signaal-

vorm of een codering daarvan vergeleken met de vormen of codes van enige referentiesignalen die typische voorbeelden zijn van de te onderscheiden klassen. De meeste werkende cijferherkennende systemen berusten op deze methode.

Door Ho en Agrawala (1968) [5] is een uitgebreid overzicht gegeven van een aantal classificatiemethoden. Het geval dat de objecten worden gerepresenteerd door kwalitatieve kenmerken is tot nu toe buiten beschouwing gebleven. Meestal wordt m.b.v. a priori kennis een zgn. beslissingsboom gemaakt, waarbij de kenmerken sequentieel worden bekeken. Er zijn methoden die automatisch m.b.v. leerobjecten een dergelijke beslissingsboom kunnen leveren. Dit is o.a. beschreven door Hunt (1966) [6].

5. Toepassingsgebieden

De meeste toepassingen van patroonherkennen worden gevonden in gebieden waar men een menselijke herkenner wil simuleren vanwege het routinematig karakter van zijn werk, of vanwege het feit dat men pas na automatisering van het herkenningsproces in staat is om complexe of massale hoeveelheden gegevens te verwerken. De volgende toepassingsgebieden vallen onder meer te onderscheiden:

- Medische en biologische toepassingen. Het stellen van diagnoses uit meetgegevens, zoals elektrocardiogrammen en elektroëncefalogrammen. Het analyseren van cellen en chromosomen ten einde afwijkingen te kunnen vaststellen.
- Automatische persoonsidentificatie aan de hand van bijv. vingerafdruk, handschrift, spraak of het gelaat.
- Communicatie tussen mens en machine met behulp van gesproken en geschreven woord. Hieronder valt o.a. het automatisch lezen van letters en cijfers, met onder meer als doel het toegankelijk maken van bestaande archieven.
- Het verwerken van luchtfoto's t.b.v. luchtkartering, weersvoorspelling, militaire toepassingen, e.d.

Naast deze gebieden zijn er nog toepassingen als het herkennen van bellenvatfoto's, het interpreteren van chemische analyses en het herkennen van diverse spectra, signaalvormen, e.d.

6. Vooruitzichten

De geschetste aanpak van het automatisch patroonherkennen komt er op neer dat zo ver als mogelijk is de menselijke kennis van de objecten wordt gebruikt. Op het punt dat deze tekort schiet, of moeilijk exact kan worden weergegeven, wordt de hulp van de statistiek of de heuristiek ingeroepen. Deze stellen echter duidelijke beperkingen met betrekking tot het aantal mogelijke parameters als metingen en kenmerken. In principe kan dit aantal wel groot zijn, maar het aantal benodigde leerobjecten stijgt dan tot onverwerkbare waarden. Het aantal mogelijke parameters wordt beperkt door het aantal leerobjecten, terwijl dit weer beperkt wordt door economische factoren en de grootte van de gebruikte rekenapparatuur.

Een en ander heeft tot gevolg dat het merendeel van de herkenningsproblemen onmogelijk geautomatiseerd kan worden wanneer men bij de opzet afziet van het inroepen van vèrgaande hulp van de menselijke herkenner. Voor problemen die in het geheel niet door de menselijke herkenner kunnen worden opgelost is, afgezien van een aantal detectieproblemen, automatisering vrijwel in het geheel niet mogelijk. Hierdoor kan men zich afvragen in hoeverre de gekozen statistische en heuristische aanpak van het herkennen wel de juiste is. Bij onderzoek in deze richting kan het van belang zijn na te gaan hoe de menselijke herkenner in het algemeen werkt, dus niet toegespitst op een specifiek probleem. De onderzoekingen van Leeuwenberg (1970) gaan in deze richting [7]. Deze sluiten aan bij de zogenaamde taalkundige benadering van patroonherkennen. Hierbij worden objecten beschreven met behulp van een 'taal' met een daarvoor gedefinieerde grammatica. Op basis van deze grammatica wordt geclassificeerd.

Literatuur

[1] BAKIS, R., HERBST, N. M. and NAGY, G.: An experimental study of machine recognition of handprinted numerals. IEEE trans. systems science and cybernetics, Vol. SSC4, pp 119 ... 132, July 1968.

[2] COOPER, P. W.: Nonsupervised learning in statistical pattern recognition. Methodologies in Pattern Recognition, edited by S. Watanabe, Academic Press, N.Y., 1969.

[3] COVER, T. M. and HART, P. E.: Nearest neighbour pattern classification. IEEE trans. information theory, Vol. IT-13, pp. 21 ... 27, January 1967.

[4] FU, K. S.: Learning control systems. IEEE trans. automatic control, Vol. AC-15, pp. 210 ... 221, April 1970.

[5] HO, Y. C. and AGRAWALA, A. K.: On pattern classification algorithms; Introduction and survey. Proc. of the IEEE, Vol. 56, no. 12, December 1968.

[6] HUNT, E. B., MARTIN, J. and STONE, PH. J.: Experiments in induction. Academic Press, N.Y., 1966.

[7] LEEUWENBERG, E. L. J.: Perceptual coding language for three-dimensional shapes and auditory patterns. Funktiepsychologisch informatie bulletin, no. 5, Psychologisch Laboratorium, Katholieke Universiteit Nijmegen, november 1970.

[8] LEVINE, M. D.: Feature extraction: a survey. Proc. of the IEEE, Vol. 57, no. 8, August 1969.

[9] NAGY, G.: State of the art in pattern recognition. Proc. of the IEEE, Vol. 56, no. 5, May 1968.

[10] NILSSON, N. J.: Learning Machines. McGraw-Hill Books, 1965.

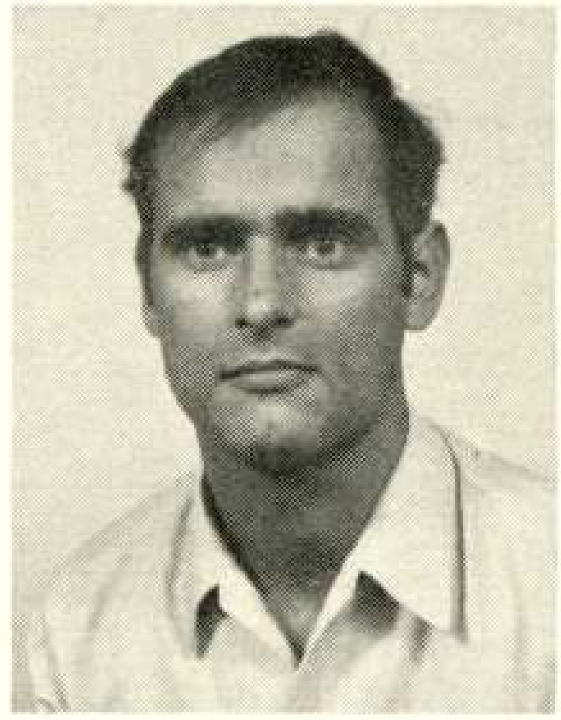
[11] PATRICK, E. A.: On a class of unsupervised estimation problems. IEEE trans. information theory, Vol. IT-14, pp. 407 ... 415, May 1968.

[12] SKLANSKY, J.: Learning systems for automatic control. IEEE trans. automatic control, Vol. AC-11, pp. 6 ... 19, January 1966.

[13] SPECHT, D. F.: Generation of polynomial discriminant functions. Technical report No. 6763-5. Stanford Electronics Lab., Stanford University, May 1966.

II. A General Coding System, Simulating the Human Classification of Visual and Auditory Patterns

by dr. E. L. J. Leeuwenberg, Psychological Laboratory, Catholic University, Nijmegen



Synopsis: A theory is presented dealing with the specification of the structural units of information that are independently processed in the human perception of three-dimensional figures (shapes) and auditory patterns. This theory implies that all types of shapes can be encoded in a formal language, which is maximally efficient and exhaustive. It is hoped that this language can give suggestions for the construction of a general pattern-recognizer.

1. Model for the use of structural information units by the human mind

One of the most baffling functions of the human perceptual system is undoubtedly the ability to record the outside world. While concerning *thinking* or *deduction*, numerous laws have already been inferred (logic) – so many indeed that it has seriously been questioned whether man in fact can think at all –, with regard to the *induction* of concepts, little to nothing is known. This last process occurs already at the very level of perception. It will immediately be evident that in perceiving two different triangles the two figures correspond with respect to the category of 'triangularity'. The question here is: how does the perceptual system process the data of the patterns, such that precisely 'triangularity' is selected as the relevant category, while in principle an infinite number of categories could crop up in the mind? To this question the reply could be: 'The perceiver counts the number of angles of each figure and then checks whether both figures have the same number of angles'. But what does the perceiver do with the two patterns of Fig. 1?

The common category is 'circularity'. Clearly, counting the angles does not lead to this conclusion. Could it be that the perceptual system tests all possible, i.e. an infinite number of rules in a fraction of a second, or will it consider merely a small number of highly potential rules which indirectly involve many others? We shall proceed from the latter suggestion and shall attempt to establish the existence of such universal coding principles.

Let us look at an example of how this investigation could be started. We will assume that a perceptual system is confronted with the two patterns of Fig. 2.

Suppose we could take a look behind the eye of the perceiver and could discover there how the two configurations are recorded in the brain, or, what is equivalent, how these figures are expressed in a new perceptual language. We assume for the moment that this language can be 'read'. We may then find as the code for Fig. 2a: 'concatenated small straight lengths of line, which consecutively form the following angles in relation to a fixed base: 0°, 90°, 0°.' As the code for Fig. 2b we may read: 'concatenated small straight lengths of line, which consecutively form the following angles in relation to a fixed base: 45°, 135°, 45°.' The first parts of the figure transcriptions are identical. We therefore understand that the perceiver can indicate the correspondence with respect to the straight lengths of line in the two figures. However, the latter part of the encoding does not show a ready-made similarity. The perceptual language we have used in our example is insufficient. Indeed a human perceiver would see immediately that the figures are almost identical, one being merely rotated somewhat relative to the other. Let us therefore see whether we can improve the language.

Suppose that in the brain of the perceiver another small perceiver (also known as *homunculus*) was busy with the as yet unrelated values: 0°, 90°, 0° and 45°, 135°, 45°, pairing them consecutively and subtracting them one from the other. Thus: $90° - 0°$ and $0° - 90°$ give $90°$ and $-90°$ respectively; $135° - 45°$ and $45° - 135°$ give $90°$ and $-90°$ respectively. By means of this subtracting operation, the two initially unrelated angle series are made equal. Thus the interpretative description of each figure in terms of differences between consecutive angles is in this case a better one than the translation in terms of angles in relation to a fixed base.

Let us, for the time being, merely conclude from this the existence of a rather general property. We saw that in our example the perceiver first registered the angles. Only after this registration, did the super-perceiver (homunculus) begin to perform his subtraction tricks. The second perceiver, the super-perceiver, thus was building on the result of the first perceiver. Could it be that this strategy only applies in the translation of this pair of figures, or would it have a more general applicability?

Let us apply the strategy to the translation of Fig. 1a. This could read: 'dots in the configuration of a circle, in which to each dot a square is attached'. To begin with, the first part of this description has to be determined; only then can the second

Paper, read before the audience during the 214th Technical Meeting of the NERG, on 29 January, 1971, in the Conference Room of the Royal Institution of Engineers, The Hague.

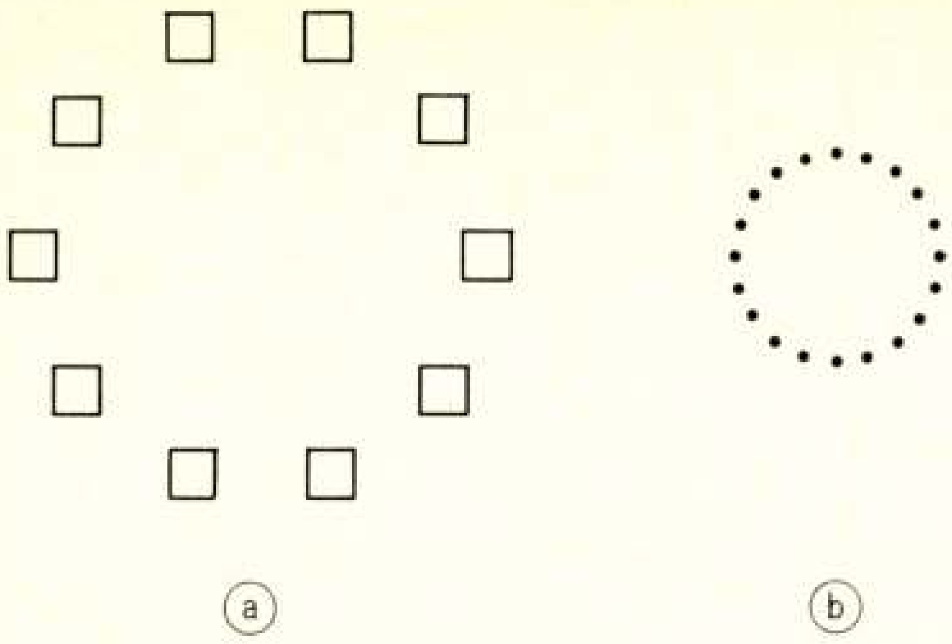


Fig. 1. Circular patterns.

part specify the first. In other words: *there is a hierarchic relationship between the two parts of the description.*

The strategy discussed above has thus been applied in the translation of the figure. The translation is moreover psychologically relevant, for the first part of the translation describes exactly the 'psychological' similarity between Figs. 1a and 1b. One might go on in this manner, devising perceptual laws and testing these. The way in which we have made a start here may be very rough indeed, but any perceptual rule conveying a broad restriction may be of use. Typical examples of such rules are provided by the '*Gestalt*' *Laws*.

We shall now analyse in more detail the pattern of Fig. 3, to enable ourselves to give a preliminary outline of the coding system which we shall discuss later on, and to illustrate more clearly the principle of perceptual hierarchy.

Let us trace how the pattern is stored by a perceiver, first at a peripheral level and subsequently at a more central level, in such a manner that, in agreement with considerations given by Attneave (1967), the recording at the latter level is less redundant than that at the more peripheral level. The peripheral recording might consist of the series of angles, formed at each dot, between its connections with the previous dot and the base axis. We then obtain, starting at point A, the following series of angles, expressed in degrees:

46,46,46,46, 23,23,23,23, 0,0,0,0, -23,-23,-23,-23, -46,-46,-46,-46, -23,-23,-23,-23, 0,0,0,0, 23,23,23,23, 46,46,46,46, 69,69,69,69, 92,92,92,92, 115,115,115,115, 138,138,138,138, 115,115,115,115, 92,92,92,92, 69,69,69,69, 46,46,46,46. (1)

Again we shall presume a homunculus, capable of transferring this series of numbers into a shorter form without, however,

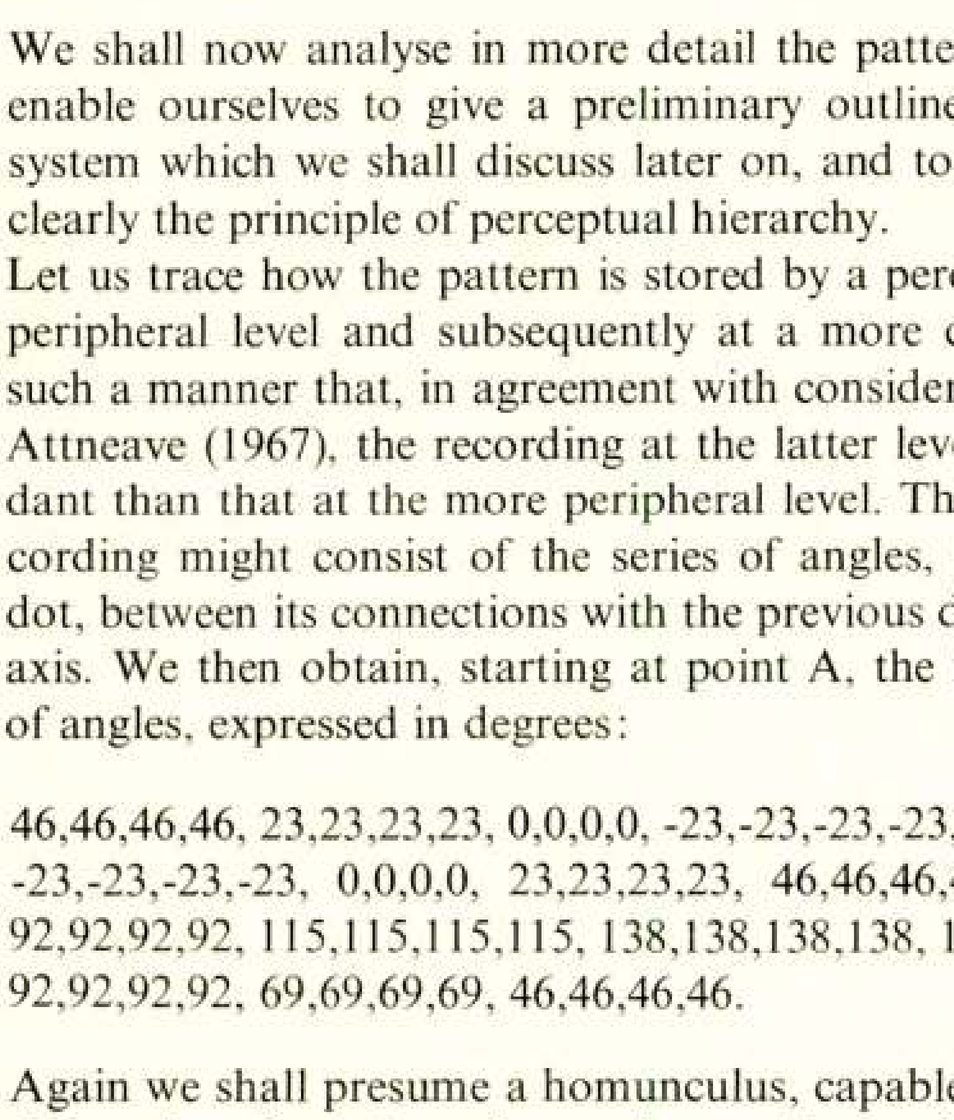


Fig. 2. Two similar patterns in differing positions.

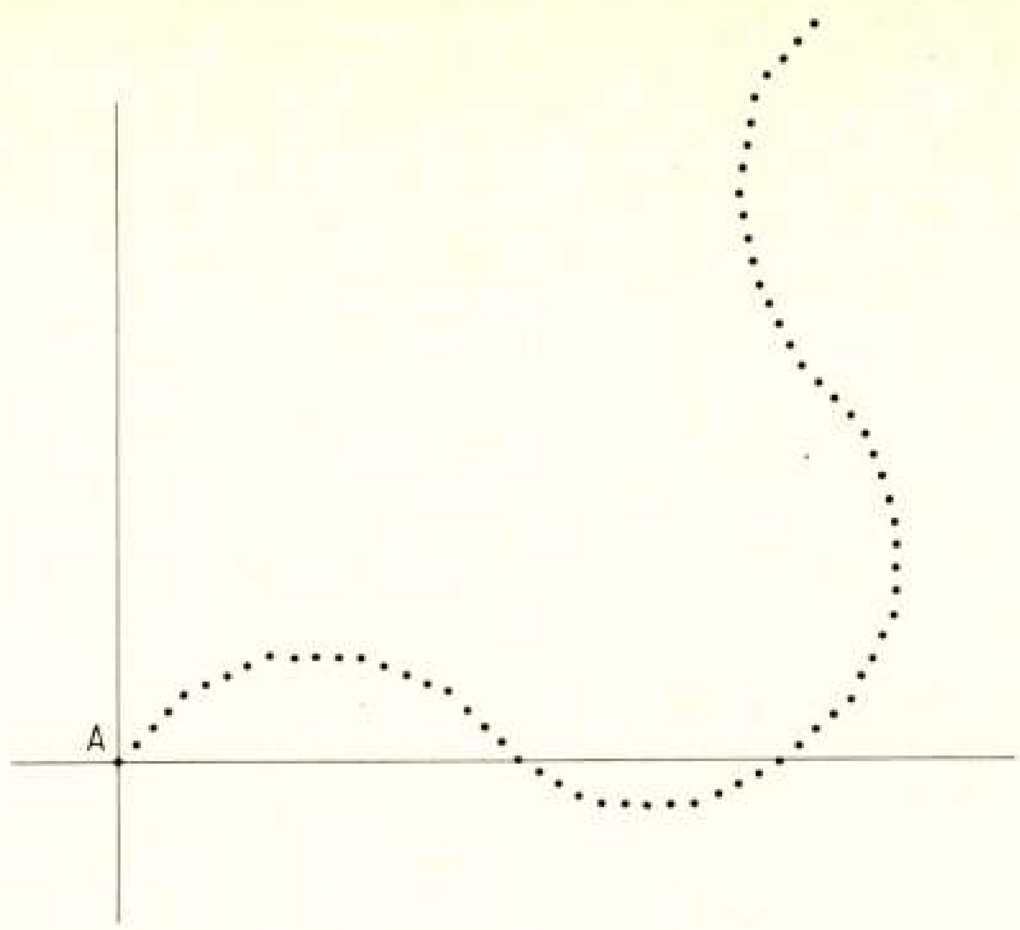


Fig. 3. Curve of dots, arranged in groups on straight lines.

causing the immediate disappearance of the first redundant recording. The quadruplicity of the angle values may be expressed as follows:

4 · (46,23,0,-23,-46,-23,0,23,46,69,92,115,138,115,92,69,46) (2)

Within the parentheses simplifications will appear, if instead of the values themselves their consecutive differences are considered, in the same way as we have seen above.

In order to show how this can be expressed symbolically, we shall now insert a minor example: 0,1,2,3,2 = / (1,1,1,-1). Between the parentheses following the / -sign, the differences are put in that have been obtained from a pairwise subtraction. Expression (2) may be translated to become:

4 · (/ (46,-23,-23,-23,-23,23,23,23,23,
23,23,23,23,-23,-23,-23,-23)) (3)

Once again quadruplicity can be noticed. It is remarked that this will be possible only after the 'translation by means of pairwise subtraction (/)' has been applied. Therefore, at the next recoding, a hierarchic transformation is performed. The result now becomes:

4 · (/ (46, 4 · (-23,23,23,-23))) (4)

Between the inner parentheses a symmetry is observed. This property of symmetry constitutes itself a simplifying perceptual principle, as has been demonstrated by Attneave (1954).

Let us agree that R $(a,b) = a,b,b,a$.
Recoding will give:

4 · (/ (46,4 · (R(-23,23)))) (5)

This might also be written as:

4 · (/ (46,4 · (R($\mp$(23))))) (6)

In this way a figure may be recorded very concisely and very efficiently. One can attempt different sequence orders in the transformations, but, after some calculation, it will appear that generally only one sequential order is the most efficient.

In this context the very pertinent question arises as to what determines the length or inefficiency of the formula. One might suggest that the greater the number of symbols in a formula,

the more inefficient it is. For instance, in the formula $3 \times (a+b)$ seven symbols may be distinguished. However, these symbols are not independent of each other. The symbol $\times$ nor the symbols ( ) have a meaning if 3 were absent, and vice versa. We therefore regard $3 \times (\ )$ as one given entity, or in other words as one 'structural information unit'. Hence, this formula contains only three information units.

As a general rule we call the symbols + and × *operations*, the ( ) symbols we call *indicators*. We attribute the name of *structural information unit* to the signs /, R, ±, and to numerical values. Thus the information of Fig. 3 amounts to 7 units. In general we ignore the information concerning the tilt of the figure, in this case 46° and the first / -transformation, so that Fig. 3 contains 5 rather than 7 units in its final code.

The coding system which we are about to discuss has in first instance bearing on two- and three-dimensional structures. The system can also be applied, as will be shown, to auditory patterns. The coding system for two-dimensional figures has been described elsewhere (Leeuwenberg 1968, 1969). Nevertheless we shall frequently return to this, because the encoding of three-dimensional structures contains simplifications analogous to those of the two-dimensional coding procedure. At first our pattern-language will make the impression of being very complicated, both because the analogies become clear only afterwards, and because we present the system of rules in a very formal way.

In presenting this model we do not wish to suggest that no other possible codes of figures do exist, such as those occurring specifically at short-term memory level, or in imagination (Moran 1969). Moreover we leave out of consideration the manner in which in a certain context representations of structures may be laden with meaning, when evoked in a human perceiver. In other words, the coding system we present is related to the ultimate cognitive representation of structures to be expected in long-term memory, in so far as these structures are free of meaning. In this coding system various perceptual properties can be abstracted, such as: 'information', 'similarity', 'hierarchy', 'regularity', and 'redundancy of figures'. By means of these quantifiable properties quantitative tests of perception, of organic cerebral disorders, and of fluid intelligence (Cattell 1968) may be developed eventually. Another effect of using this coding system could be that it may yield suggestions for models of pattern recognition.

In our discussion, an attempt will be made to substantiate the choice of coding rules. In the next section these coding rules will be set out. Subsequently, we shall consider some examples for illustrative purposes.

2. Coding Rules

2.1. *Structural Information Units.*

1. n – 'number', as in: $a,a,a, = 3 \cdot (a)$. This information may specify any size.
2. / – 'integration': / (3,2,5) = 0, 0+3, 0+3+2, 0+3+2+5 = 0, 3, 5, 10.
3. R – 'reversal': R[3,2,5] = [3,2,5], [3,2,5].
4. ± – 'left-right variation': ±(90°) = 90°, −90°.

2.2. *Operators.*

1. + – addition: $5+3 = 8$.
2. × – multiplication: $5 \times 3 = 15$.
3. · – iteration: $3 \cdot (a) = a,a,a$.

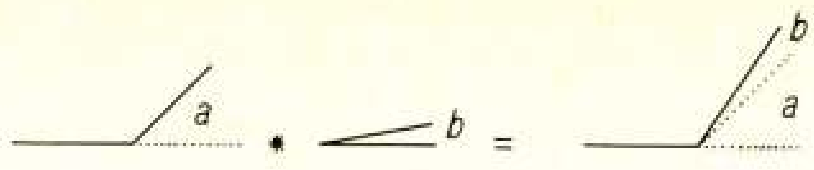


Fig. 4. Addition of two angles.

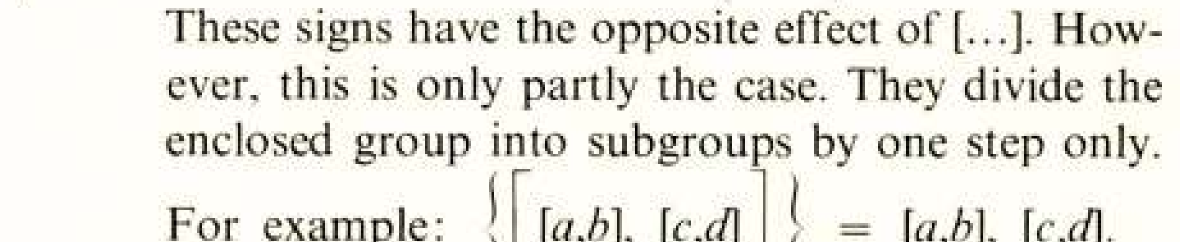
4. ∗ – angle addition: In analogy to the addition of angles in the multiplication of vectors; unlike vector multiplication, this operation does not affect length (see Fig. 4).
5. ⊛ – addition of coincident angles; see Fig. 5.

This operation holds only if two figures meet each other (here in point P). The operation is used in order to maintain certain patterns – like in Fig. 5 |10,10,10,10,10,10| – invariant under topological transformations such as the 'curved surface' |0,0,0,-45,45,45|. We shall give examples of this further on, when discussing Figs. 10 g, j, l.

It will be made clear later, why these numbers are usually enclosed by 'the 'absolute signs' |...|.

2.3. *Indicators.*

1. (...) – *continuation:* $(a) = a,a,a,$ ...... etc. This continuation holds until the sequence meets 'something', e.g. a line or a surface.
2. [...] – *chunking:* $2 \cdot [a,b] = a,b,a,b$. The sequence within these signs is treated as one unbroken unit in every context.
3. (...) – *border signs:* These signs are rather meaningless and often redundant. For example, $2 \cdot (a,b) = a,a,b,b$. The enclosed values do have a common fate, but function separately. Furthermore, these signs enable the applicability of the rule that in the reconstruction of a formula the most deeply embedded values are solved first.
4. {...} – *breakdown into elements:* {[a,b]} = a,b. These signs have the opposite effect of [...]. However, this is only partly the case. They divide the enclosed group into subgroups by one step only. For example: {[[a,b], [c,d]]} = [a,b], [c,d].
5. {...} – *take sign:* the functioning of an enclosed value continues until it has been combined with all the



{|0,0,0,45,45,45|} ⊛ {|10,10,10,10,10,10|} = |10,10,10,10,55,55|
= |0, 0, 0,45,45,45|

Fig. 5. Addition of angles over shifted junctures.

unenclosed adjacent values. $\{2,3\} \cdot (a,b) = 2 \cdot (a), 2 \cdot (b), 3 \cdot (a), 3 \cdot (b)$. The normal situation is: $(2,3) \cdot (a,b) = 2 \cdot (a), 3 \cdot (b)$.

6. $\{\ldots\}$ – *complete breakdown into elements:* These signs represent the total opposite of the [...] signs. They break the strongest chunk into its utmost smallest elements. For example: $\{[[a,b], [c,d]]\} = a,b,c,d$.

7. | ... | – *absolute signs:* The enclosed values refer to angles determining sequentially connected straight lengths of line. These angles are, however, related to one reference base-axis. When angles are not enclosed by these signs, then every value refers to the angle between a piece of line and the continuation of the previous line. They form 'relative angles', for example: |17°, 33°| = 17°, 16° (see Fig. 6).

8. ... || ... – *independence of angles:* (*a*,*b*) || (*c*,*d*), meaning that *a* and *b* do not form the reference axis for the angles *c* and *d*. (In this case *c* and *d* are dependent on *a* and *b* in another aspect.)

When *a* and *b* in their turn are dependent on *p* and *v*, while *c* and *d* are also angularly independent of *p* and *v*, this can be indicated by ((*p*,*v*)(*a*,*b*)) |||| (*c*,*d*), where |||| is substituted for ||.

9. ⟨ ... ⟩ – *outerproduct:* 14°, 67° ⟨45°⟩ (Fig.7). In this series the last angle (45°) is measured in the plane, determined by the outerproduct of the reference axis of 14° and 67° and by the reference axis of the angle of 45° (which is the second leg of the angle of 67°). This plane is perpendicular to the plane in which the angle of 67° is positioned, and contains the perpendicular that can be erected on the latter plane. This perpendicular matches to the plane of the angles of 14° and 67° in accordance with the right-hand rule when turning the reference axis, determined by the angle of 14°, towards the reference axis, determined by the angle of 67°. The positive direction for the angle of 45° is given by turning its second leg around its juncture from the reference axis towards the perpendicular, as shown in Fig. 7.

In cases where there are more included angles, as in: 14°, 67°, ⟨−90°, 45°⟩, Fig. 8, every included angle is also an 'outerproduct' of the reference axis of the last two angles. Hence, 45° is an 'outerproduct' of 67° and −90°. However, the right-hand rule must be applied to the rotation determined by 67°, +90°, in order to avoid a change in the structure represented by following elements in the formula, if the second angle (in this case −90°) of the couple determining the outerproduct, passes from a positive to a negative value, or vice versa. In general, the *right-hand* rule has to be applied to the rotation of the reference axis of two successive angles, if the second angle is *positive*, while the *left-hand* rule must be applied if this second angle is *negative*, to make the succeeding structure invariant with respect to the sign of the second angle of an outerproduct-couple.

In the case {14°, 67°} ∗ {⟨|90°, 45°|⟩} the 45° is not an 'outerproduct' determined by the last two angles 67°, 90°, but determined by the pairs of angles at the left side of the ∗-sign (in this case 14°, 67°). This exception only is found when the operation ∗ is present. Later we shall give an explanatory example (see Fig. 10 l).

10. —— – *vanishing sign:* (see Fig. 9).

All the values function normally as existing elements, but they are invisible. An exception to this rule is the case in which there would be no difference between the situation with, and without,

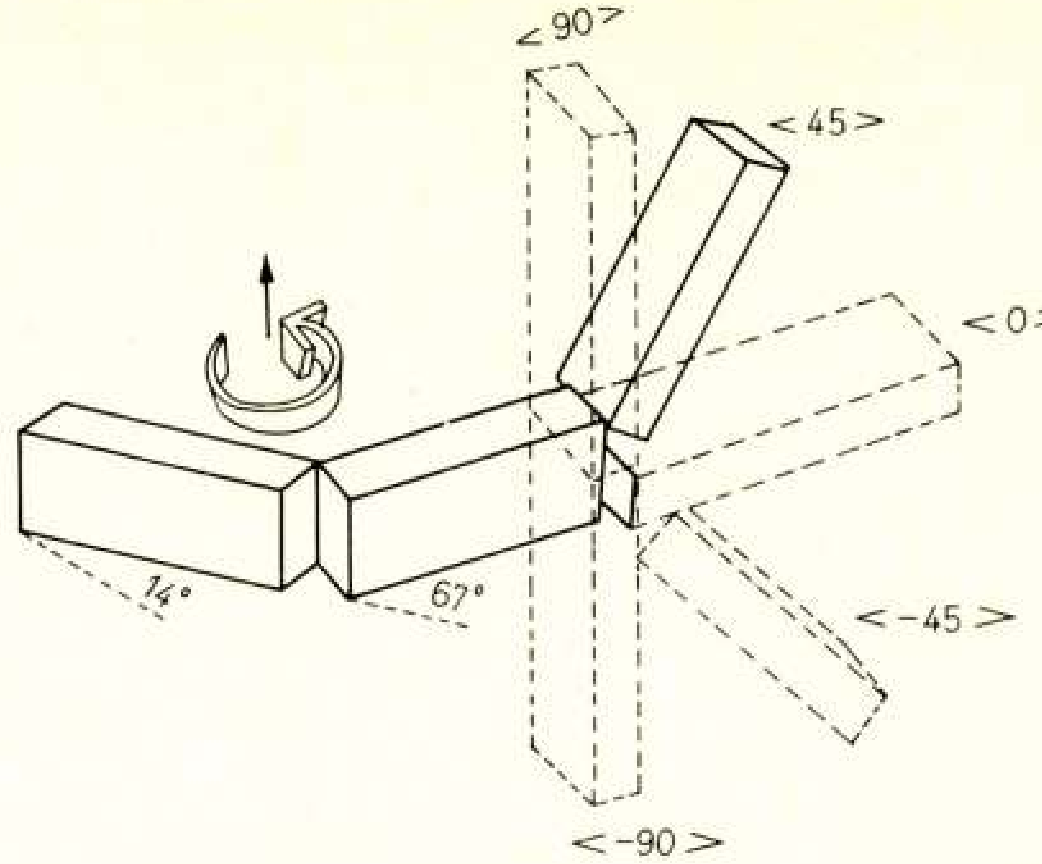


Fig. 7. Outerproduct of positive angles.

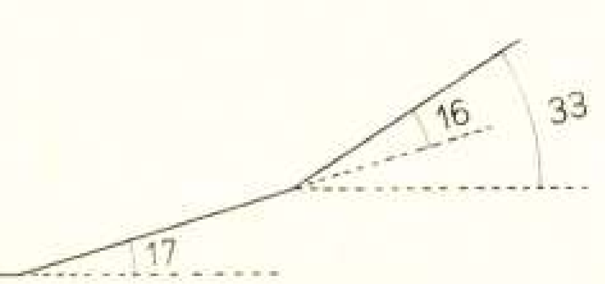


Fig. 6. Interrelation of absolute and relative angles.

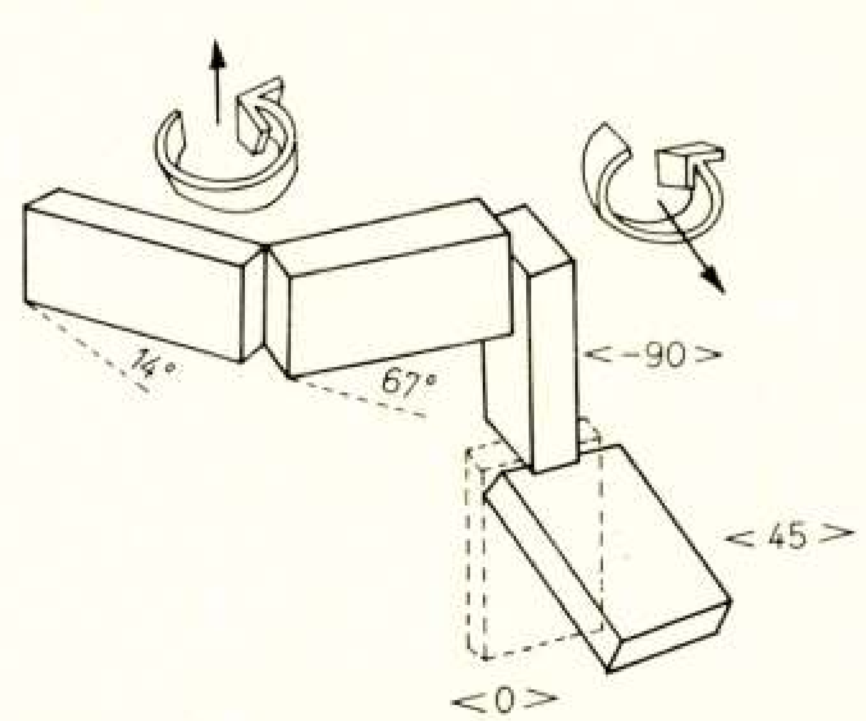


Fig. 8. Outerproduct, when the second angle is negative.

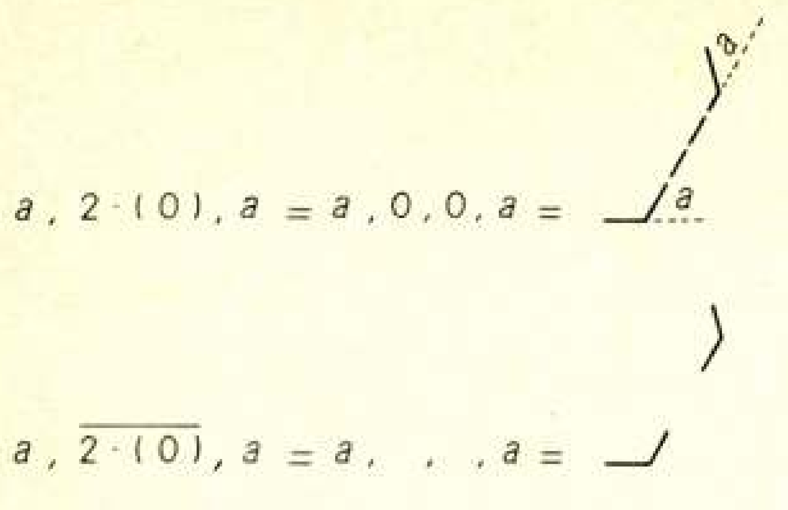


Fig. 9. The sense of the vanishing sign.

signs; then the values do not function normally (see description Fig. 10 c, d, h).

3. Combinatory and Implicit Rules

Combination:

$(a,b)(c,d) = a, c, b, d$
$(a,b)[c,d] = a, [c, d], b, [c, d]$
$[a,b](c,d) = [a, b], c, [a, b], d$

For Operations:

$(a,b) + (c,d) \quad = a+c, b+d$
$(a,b) + [c,d] \quad = a+[c, d], b+[c, d]$
$a,b \quad + (c,d) \quad = [a+c, a+d], [b+c, b+d]$

These three hold for all the operations $(+, \times, \cdot, *)$. For + and * holds: $a+[b, c] = [a+b, c]$. For iteration: $2 \cdot (a\ b) = a,a,b,b$ and $2 \cdot [a\ b] = [a,b], [a,b]$.

For Onesided Iteration:

(3); $(a,b)(c,d) = a,b,a,c,b,a,b,d$

Iteration is wanted only for the first group, indicated by the point in the sign ;. This means: take three elements from (a,b), then one from (c,d), continue with three elements from (a,b) and one from (c,d), etc.

(2): $(a,b)(c,d,e) =$ a,c,d,b,e,c,a,d,e,b,c,d,a,e,c,b,d,e

In this second example iteration is wanted for the second group, indicated by the point in the sign :. This means: take one element from (a,b), then two elements from (c,d,e); continue with one element from (a,b) and two elements from (c,d,e), etc. Likewise we have: $(1 \cdot 2)$: $(a)\ (b) = a, b, a, b, b$.

The same can be applied to $\pm$, viz.:
$3 : (\pm) = + + + -$
$3 : (\pm) = + - - -$
$3 \cdot (\pm) = + + + - - -$
$3 \cdot [\pm] = [+ -][+ -][+ -]$

Serial Structure: $a, 4 \cdot (0), a, 4 \cdot (0)$, etc. $= \{a, [4 \cdot (0)]\}$ (see Fig. 10a).

At nodes, from which different branches arise, we use the parallel structure representation; in other configurations the normal serial structure is used.

Parallel Structure:

$a, 4 \cdot (0)$
$a, 4 \cdot (0)$
$a, 4 \cdot (0) \quad = \widehat{a, 4 \cdot (0)} \qquad (10b)$
$a, 4 \cdot (0)$
etc. (see Fig. 10b).

The upper angles form the reference bases for the lower angles and have common starting points.

4. Examples and Analogies

With the aid of some examples, we shall now attempt to make more clear how visual structures of two and three dimensions as well as auditive structures may be encoded by means of these coding rules. Generally we shall in a first stage express visual patterns in terms of relative angles (see section **2**, Indicator 7 and Fig. 6). Further, let us agree that the first ten letters of the alphabet be used to symbolize angles, while the remaining letters represent numbers. Accordingly, Fig. 10a may be encoded, as we saw above, in the following manner:

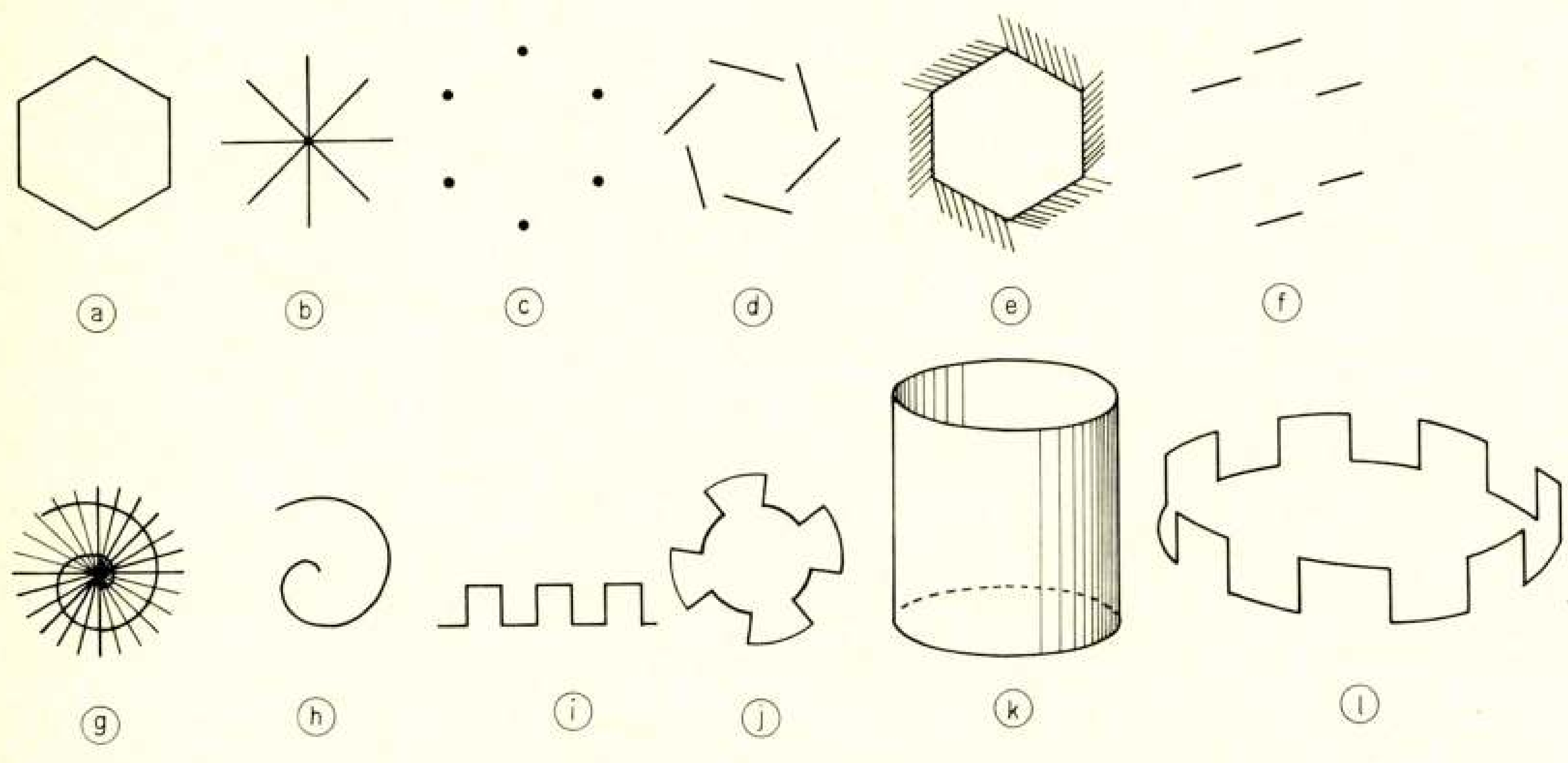


Fig. 10. A series of figures of increasing complexity.

$\left\{a,[n\cdot(0)]\right\}$. If n is sufficiently small, this polygon approaches a circle. Thus both configurations contain two information units. Fig. 10b may be described by: $\overbrace{a},[\overbrace{n\cdot(0)}]$ .Just as in this case the $\{\ \}$ symbol may be used for serial as well as for parallel formula-structures, so all combinational rules may be used both for the two-dimensional, serial, horizontal formula-structures, and for the parallel, vertical formula-structures which are primarily suitable for three-dimensional figures. The two-dimensional parallel structure of the formula of Fig. 10e is an example of this. Its code reads:

$$\begin{array}{c}\overbrace{\left\{a,[n\cdot(0)]\right\}}^{\circ} \\ \underbrace{\hphantom{xx}}_{\circ} \\ \boxed{b,\ p\cdot(0)}\end{array} \qquad (10e)$$

This means: take an element (angle) of the polygon $\left\{a,[n\cdot(0)]\right\}$ and attach to it a straight line $p\cdot(0)$ of length p. This straight line is linked at an angle b to this element of the polygon. As the combination rule (given by the brackets) dictates, one must now return to the next element (angle) of the polygon $\left\{a,[n\cdot(0)]\right\}$. This element should now be attached, not to the straight line $p\cdot(0)$ just mentioned, but, as the vertical structure of the formula indicates, to the element of the polygon mentioned first. Subsequently attach to this, again, a straight line $p\cdot(0)$ at angle b; and so on in the same manner.

An example in which a vanishing sign finds expression is Fig. 10c, whose code reads:

$$\left\{a,[\overline{n\cdot(0)}]\right\} \qquad (10c)$$

Analogous to Fig. 10c, Fig. 10d may be written as:

$$\begin{array}{c}\overbrace{\left\{a,[\overline{n\cdot(0)}]\right\}}^{\circ} \\ \underbrace{\hphantom{xx}}_{\circ} \\ \boxed{\overline{b,\ p\cdot(0)}}\end{array} \qquad (10d)$$

This already constitutes an exceptional case, as not only $n\cdot(0)$ itself disappears, but also its further function, this being revealed by the fact that were $n\cdot(0)$ has disappeared, $p\cdot(0)$ has also disappeared. If, in this case, this had not been so, the formula would have also applied to Fig. 10e. The vanishing sign would not have had any meaning, in contrast to the case of Fig. 10h.

First we shall show the meaning of the 'independence of angles' indicator No. 8. Fig. 10f may be coded as:

$$\begin{array}{c}\overbrace{\left\{a,[\overline{n\cdot(0)}]\right\}}^{\circ} \\ \underbrace{\hphantom{xx}}_{\circ} \\ \equiv \\ \boxed{\overline{b,\ p\cdot(0)}}\end{array} \qquad (10f)$$

Angle b is indeed linked to the elements (angles) of the polygon $\left\{a,\ [\overline{n\cdot(0)}]\right\}$; however, it is not dependent on the angles a, but on the 'absolute' base in relation to which the a-angles were referred in the first instance.

The following examples deal with the very complicated use of operation 5 (Section 2.2.). The formula of Fig. 10g is made up of two parts: viz. that for a star and that for a straight line. However, the line in the figure is not in fact straight, for at every point $\left(\{\circ\}\ \{\circ\}\right)$ where it intersects $(\cap)$ the radials of the star, a constant absolute angle $(|70^\circ|)$ is made with these radials ($*$). In this way the line is transformed into a spiral by virtue of its relation to the star. The representation of Fig. 10g is:

$$\circ\left\{\overbrace{a},[\overbrace{n\cdot(0)}]\right\}\circ \circledast \circ\left\{|k\cdot(70)|\right\}\circ \qquad (10g)$$

and the formula for Fig. 10h, where the star 'vanishes' from the spiral is:

$$\left\{\overline{\overbrace{a},\ \overbrace{n\cdot(0)}}\right\}\circledast\left\{|k\cdot(70)|\right\}\circ \qquad (10h)$$

The reason why the angles on the right-hand side of the formula are absolute is because they form a constant 70° with the absolute reference axes formed by the radials of the star which is described in the left-hand part of the formula (cf. Fig. 10b). In fact, the left-hand pattern must first be constructed before the right-hand pattern can be superimposed on it. Once the star pattern has transferred its directional function on the line $|k\cdot(70)|$ the star as such can vanish.

A similar process may be seen in the relation between Fig. 10i and Fig 10j. Fig. 10i, represented by $m\cdot\{|(0,\ 180)(90)|\}$, may be recognized in the formula for Fig. 10j. This is in agreement with the fact that pattern 10i can easily be recognized perceptually in pattern 10j. Fig. 10j is represented by:

$$\left\{\overline{\overbrace{a},\{0\}}\right\}\circ\circledast\left\{\overline{t\cdot(0)},\ m\cdot\circ\left\{|(0,\ 180)(90)|\right\}\circ\right\}\circ \qquad (10j)$$

The left-hand part again describes a star with infinitely long radials. Even though this star pattern vanishes, it still has its effect on the figure which corresponds to the right-hand part of the formula. This part first follows one of the radials of the star for a certain length $t\cdot(0)$ which later vanishes and then the castellated structure $m\cdot\{|0,90,180,90,0,90,180,90,0,\text{ etc.}|\}$ is attached, in such a way that it is turned by the invisible star, which provides a new base plane.

The same procedure holds for the representation of pattern 10 l, in which also pattern 10i is present, invariant under the topological surface transformation which is described in pattern 10k and represented by:

$$\begin{array}{c}\overbrace{\left\{a,\ n\cdot(0)\right\}}^{\circ} \\ \underbrace{\hphantom{xx}}_{\circ} \\ * \\ \boxed{<90>,m\cdot(0)}\end{array} \qquad (10\,k)$$

The representation of pattern 10 l is:

$$\circ\left\{\begin{array}{c}\overbrace{\left\{a,[n\cdot(0)]\right\}}^{\sim} \\ \underbrace{\hphantom{xx}} \\ * \\ \boxed{<90,0>\ 0}\end{array}\right\}\circ\circledast\left\{<m\cdot\left\{|(0,\ 180)(90)|\right\}>\right\}\circ \qquad (10\,l)$$

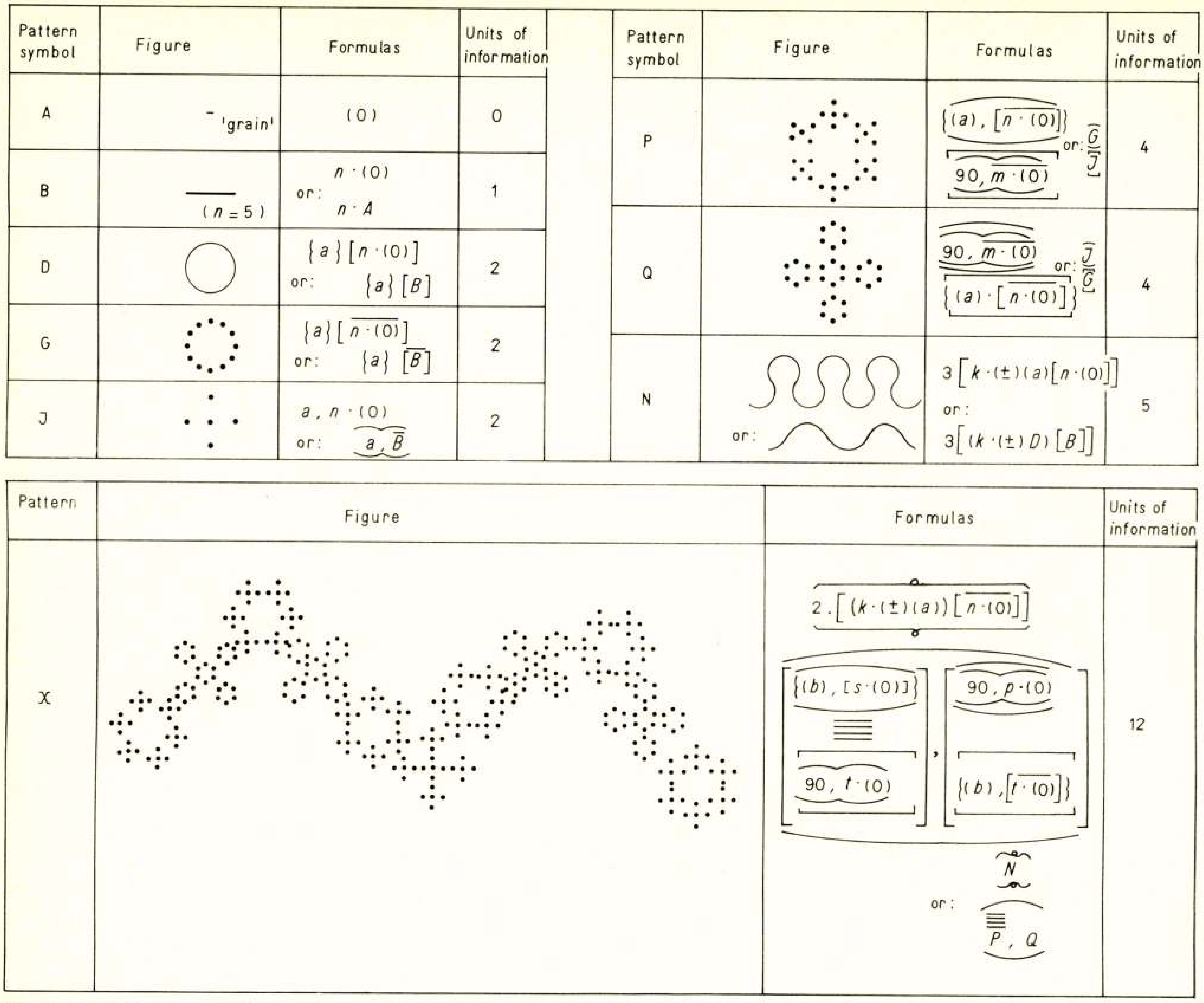


Fig. 11. Two-dimensional ribbon pattern.

Pattern 10l is composed of 10k and 10i. Pattern 10k is represented by the left-hand and 10i by the right-hand part of formula 10l. Because of the combination of the ✻ and ⟨ ⟩ signs, the angles at the right-hand side form an 'outer-product' with the surfaces, constituted by the angles of the left-hand side figure. However, this 'outerproduct' occurs only between those right-hand side angles which *coincide* with left-hand side angles. In order to specify these angles, the left-hand side figure has to be constructed first and then the right-hand side figure must be superposed on this figure step by step. So it is not possible to construct the shape of structure 10l directly from the two sub-structures 10k and 10i.

The proposed formulation of Fig. 10l is analogous to that for the construction of a drawing 10i on a curved surface. The usual way to construct Fig. 10l is to draw 10i on a sheet and then to bend the sheet in a circular way. The 10i drawing and the circular bending are the two constituents of the formula.

As another complex configuration pattern X is analysed in Fig. 11. It is a composition of the sub-structures N, P and Q. The structures P and Q are fitted into the sinus superstructure N. They contain both cruciform (structure J) and circular (structure G) patterns. Pattern X as a whole can also be considered as being built up from the fine structures A, B and D. In order to show that in pattern X the cruciform patterns are not only angularly independent of the circular structure, but also of the sinus superstructure, indicator 8 (Section 2.2.) is used there twice in pattern *P*.

In general, subprocesses in the construction of a shape correspond to the units in the representation of the shape. The operation ✻ allows also the formulation of two- and three-dimensional structures superimposed on three-dimensional shapes (see Fig. 12).

We have become acquainted with an example in which a combination rule has been applied vertically. Now we shall introduce an example in which many combination rules (also vertically applicable) are applied horizontally. For a change, we shall employ an *auditory* pattern as our example, viz. the following series of values corresponding to the pitch of tones: 2,6,-2,7,3,6,-3,7,0,8,-1,6,2,7,-4,6,4,7,-5,8,5,6,-3,7,4,8,-4,6,3,7,-2,6,1,7,-2,8,2,6,-3,7,3,6,-0,7,4,8,-4,6,5,7,-5,6,0,7,-1,8,2,6,-3,7,3,6, etc. (S).

From it we arrive by successive steps of encoding at (see Table 1):

Table 1. Encoding steps and resultant coded representation of the series (S).

{(2,-2,3,-3,0,-1,2,-4,4,-5,5,-3,3,-3,4,4,3,-2,1,-2,2,-3,3,-0,4,-4,5,-5,0,-1,2, etc.)	(6,7,6,7,8) }
{(2,-2,3,-3,0,-1,2,-4,4,-5,5,-3,3,-3,4,4,3,-2,1,-2,2,-3,3,-0,4,-4,5,-5,0,-1,2, etc.)	·{2 · [6,7],8}·}
{(±(2,2,3,3,0,1,2,4,4,5,5,3,3,3,4,4,3,2,1,2,2,3,3,0,4,4,5,5,0,1,2,3,3,4,4, etc.)	·{2 · [6,7],8}·}
{(±((3,1):(2,2,3,3,4,4,5,5,3,3,4,4,2,2,3,3, etc.)(0,1,2,3,3,2,1,0)))	·{2 · [6,7],8}·}
{(±((3,1):(2,2,3,3,4,4,5,5,3,3,4,4,2,2,3,3, etc.)·{R[0,1,2,3]}·))	·{2 · [6,7],8}·}
{(±((3,1):(2,2,3,3,4,4,5,5,3,3,4,4,2,2,3,3, etc.)·{R[⌊ (1,1,1)]}·))	·{2 · [6,7],8}·}
{(±((3,1):(2,2,3,3,4,4,5,5,3,3,4,4,2,2,3,3, etc.)·{R[⌊ (3 · (1))]}·))	·{2 · [6,7],8}·}
{(±((3,1):·{ (1,3,2,) +[1,1,2,2]}·{R[⌊ (3 · (1))]}·))	·{2 · [6,7],8}·}

Finally:

{(±((3,1):·{ (1,3,2) +[2 · (1,2)]}·{R[⌊ (3 · (1))]})) ·{2 · [6,7],8}·}

This is the shortest structural description of the above pattern (which, in fact continues much farther than is given here) without any recurrence. The number of different information units is 17. They are pointed out by arrows beneath the final formula.

A further analogy between two- and three-dimensional structures is this: for the description of two-dimensional structures, the relative angle forms the basic element. This relative angle indicates only how a straight line is determined if the previous straight line is known or determined. Similarly, for the third dimension, a straight line can be determined via the 'outer-product' only if the two preceding straight lines, whose sequential order is also given, are known or determined. In general, therefore, it applies that not absolute data, but rather the elementary internal relationships in this system constitute the basic data for coding.

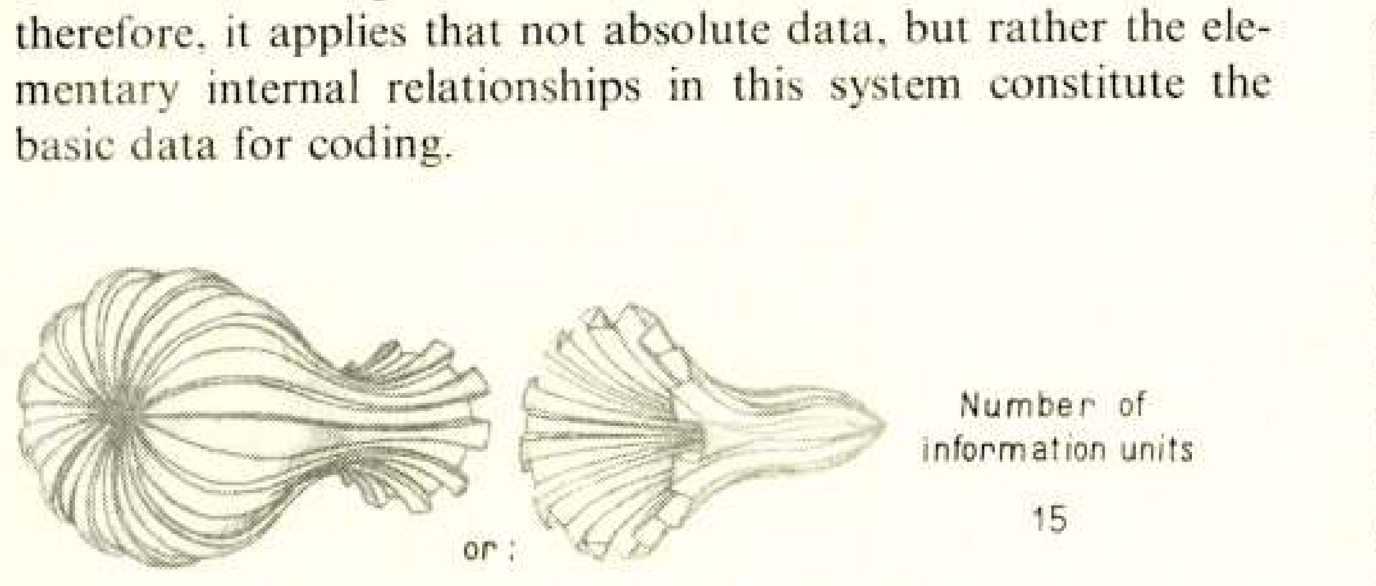


Fig. 12. Three-dimensional figure, on the surface of which a three-dimensional pattern is introduced.

Discussion

Up till now only the formal coding language has been presented, but the arguments that make clear that this system is analogous to the human perceptive system have not yet been discussed. However, elsewhere (in the American Journal of Psychology, in press, Leeuwenberg) is demonstrated how from this system predictions about the human perception of patterns are made and tested.

With respect to the problem of how to automize pattern recognition processes, this coding language can give suggestions for the building of a general pattern recognizer. It seems not so difficult to me to construct a program which enables a computer to draw the figures, given the formula; however, it is more troublesome to propose a program that would enable an automaton to develop the most economical formula, given the figure. Such an automaton may be thought of as the initial stage in the building of automatons which can handle specific sets of stimuli, e.g. characters or numerals. In the latter case, new relevant features may be evoked by the actual limitations of the set of stimuli and the specificity of the task facing the computer on the basis of different frequencies occurring within the set of stimuli.

Acknowledgement

With due thanks, the present author acknowledges the fact that the research reported on was supported by the Niels Stensen Foundation and the Netherlands Organization for the Advancement of Pure Scientific Research (Z.W.O.).

References

[1] ATTNEAVE, F.: Some informational aspects of visual perception. Psych. Rev., 1954, 61, 183 ... 193.

[2] ATTNEAVE, F.: Criteria for a tenable theory of form perception,

in 'Models for the perception of Speech and Visual Form'. Ed.: Weiant Wathen Dunn, M.I.T. Press, London, England, 1967.

[3] HANDEL, S., GARNER, W. R.: The structure of visual pattern associates and pattern goodness. Percept. Psychophys., 1966, 1, 33 ... 38.

[4] HOCHBERG, J., MCALLISTER, E.: A quantitative approach to figural 'goodness'. J. Exp. Psych., 1953, 46, 361 ... 364.

[5] LEEUWENBERG, E.: 'Structural Information of Visual Patterns'. Mouton Press, The Hague, 1968.

[6] LEEUWENBERG, E.: Quantitative specification of information in sequential patterns. Psych. Rev., 1969, 76, 216 ... 220.

[7] MACKAY, D. M.: Quantal aspects of scientific information. Phil. Mag., 1950, 41, 289 ... 301.

[8] MORAN, T.: The Grammar of visual imagery. Unpubl. Manuscript. Carnegie-Mellon.

[9] SIMON, H. A., SUMMER, R. K.: Pattern in music. Chapter 8 in 'Human Judgment'. Ed. B. Kleinmutz, publ. John Wiley & Sons, Inc., 1968.

[10] SHANNON, C. E., WEAVER, W.: 'The Mathematical Theory of Communication'. Urbana, Univ. of Illinois Press, 1949.

[11] VITZ, P. C.: Information, run structure and binary pattern complexity. Perc. Psychoph., 1968, 3, 275 ... 280.

III. Stand van de techniek met betrekking tot het automatisch lezen van met de hand geschreven cijfers

door A. A. Spanjersberg, Dr. Neher-Laboratorium, PTT, Leidschendam



Synopsis: *State of the Art in Automatic Reading of Handwritten Digits.*
A description is given of the principles of some of the reading machines, which have the possibility of reading stylized handwritten digits. The methods of scanning and recognition applied in these machines are discussed. Attention is paid to the recent development of optical reading terminals. Practical results are given, obtained in automatic reading of non-stylized digits on letters by applying an optical reading-sorting machine. A project of present interest in the field of automatic reading is mentioned.

It is recommended that a *reference set* of handwritten digits be composed. Some considerations are given which could be useful at the composition of such a reference set.

1. Inleiding

Bij het toepassen van elektronische informatieverwerkende machines voor administratieve doeleinden, bleek er behoefte te zijn aan methoden om visueel leesbare informatie direct in die machines in te voeren. Bij de studie van dit probleem werd men geconfronteerd met de verschillen tussen het menselijke perceptie- en interpretatievermogen en de mogelijkheden die machines op dit gebied bieden.

Omstreeks 1955 werd op velerlei plaatsen, verspreid over de gehele wereld, de studie van het machinale lezen aangevat. Er is sindsdien een aantal publikaties over dit onderwerp verschenen en er zijn ook veel praktische toepassingen gerealiseerd, wat betreft het lezen van informatie aangeboden in de vorm van *gedrukte* tekens of in die van machineschrift. De wenselijkheid om te kunnen beschikken over leesmachines voor met de *hand geschreven* schrift werd eveneens in een vroeg stadium van de automatisering bij administratieve processen onderkend.

Bij het bestuderen van de mogelijkheden tot automatisering van de Postcheque- en Girodienst werd gezocht naar een mogelijkheid om de geschreven numerieke informatie van rekeningnummer en bedrag automatisch in het centrale boekingsproces te kunnen invoeren. In de beginjaren van de automatisering, rond 1957, waren de technische mogelijkheden hiervoor belangrijk minder dan thans. Er werd gedacht aan documenten met zgn. merkroosters, waarvan in fig. 1 een voorbeeld is aangegeven. Hiermee zijn proeven genomen, waarbij ook rekeninghouders werden ingeschakeld.

In technisch opzicht leverde de verwerking van dit soort documenten geen problemen op. Veel werd echter gevraagd van het aanpassingsvermogen van de mens. Het is dan ook uit ergonomische overwegingen, dat men van dit soort invulroosters, althans bij de postgiro, voorlopig is afgestapt.

Bij de industrie werden soortgelijke systemen ontwikkeld. Voor een aantal toepassingen worden ze nog gebruikt. Gezien de ergonomische moeilijkheden bij het invullen en teruglezen van de merkrooster-documenten blijft de aandacht toch gericht op het machinaal herkennen van *geschreven* cijfers.

2. Industriële ontwikkelingen

Voor zover mij bekend, was de eerste die een redelijk werkend model van een leesmachine voor met de hand geschreven cijfers construeerde Dr. Sprick, verbonden aan het IBM-laboratorium in Böblingen (W.-Duitsland). Zijn ideeën zijn later in Amerika

GULDENS			CENTEN		REKENINGNUMMER						
●	0	●	0	0	0	0	0	0	0	0	0
1	1	1	●	1	1	1	1	1	1	1	1
2	●	2	2	2	2	●	2	2	2	2	2
3	3	3	3	3	3	3	3	3	●	3	3
4	4	4	4	4	4	4	4	●	4	4	4
5	5	5	5	5	●	5	5	5	5	5	5
6	6	6	6	6	6	6	●	6	6	6	6
7	7	7	7	7	7	7	7	7	7	●	7
8	8	8	8	8	8	8	8	8	8	8	●
9	9	9	9	●	9	9	9	9	9	9	9

Fig. 1. Voorbeeld van een 'merkrooster'.

nader uitgewerkt, o.a. door Greanias, en dit heeft geleid tot de machine die door IBM onder het typenummer 1287 aan de markt wordt gebracht. Er wordt bij deze machine een bijzondere aftastmethode toegepast, een variant op de zgn. 'flying-spot scanner'. Bij de 'flying-spot' aftastmethode (fig. 2) wordt op het scherm van een kathodestraalbuis een lichtstip gevormd. Door middel van een lens wordt de lichtstip op het af te tasten document afgebeeld. Een deel van de door het papier gereflecteerde hoeveelheid licht wordt met behulp van een andere lens op de kathode van een fotomultiplicatorbuis geworpen. De positie van de lichtstip wordt bestuurd door middel van de afbuigstroom in de deflectiespoelen van de kathodestraalbuis. In veel gevallen wordt een lijnvormige aftasting toegepast. De waarde van het video-signaal is evenredig met de plaatselijke reflectiecoëfficiënt op het document. In de genoemde leesmachine type 1287 vindt een contouraftasting plaats. Hierbij volgt de lichtstip de contouren van een af te tasten teken. De lichtstip maakt een gecompliceerde cirkelvormige beweging (fig. 3). Deze ontstaat o.m. wanneer een sinusvormig met de tijd verlopende stroom door de horizontale afbuigspoel en een cosinusvormig met de tijd verlopende stroom door de verticale afbuigspoel wordt gevoerd. Voorts wordt de versterking van de versterkers die deze afbuigstromen leveren gedurende het aftasten van zwarte beelddelen gehalveerd. Daardoor ontstaat het volgeffect langs de lijnen, zoals afgebeeld in fig. 4. Bij een richtingsverandering van 180°, zoals ook zou optreden bij een

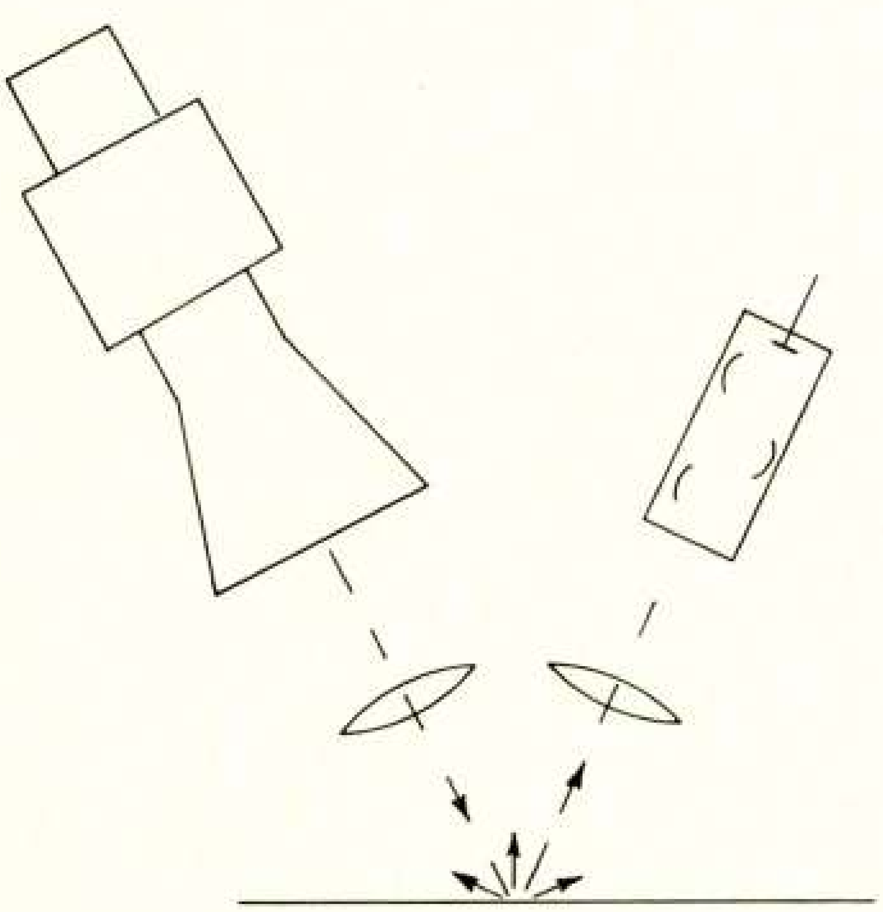


Fig. 2. Principe van de flying-spot aftasting.

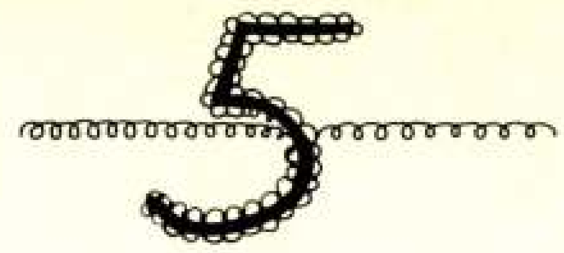


Fig. 3. De contouraftasting van een teken.

kleine lijnonderbreking, wordt getracht of het mogelijk is met een grotere amplitude van de afbuiging de onderbreking te overbruggen. In fig. 5 is deze werkwijze aangegeven. De beschreven aftastprocedure vindt voor elk teken tweemaal plaats.

De eerste maal wordt de grootte van het cijfer bepaald en wordt een denkbeeldige rechthoek gevormd die om het cijfer past. De rechthoek wordt verdeeld in een matrix van 4 × 5 eenheden (fig. 6). In elk der deelvakjes wordt de lokale richting van de lijn in eenheden van 30° bepaald. Men onderscheidt dus 12 mogelijkheden. Deze richtingen komen overeen met de kompasrichtingen. Er zijn 20 deelvakjes. De informatie-inhoud van een cijferafbeelding kan dus beschreven worden met 9 bits, nl. 5 voor de positie van het vakje en 4 voor de lokale lijnrichting.

Gedurende de wereldtentoonstelling, die in 1964 in New York werd gehouden, werd deze leesmachine gedemonstreerd. Bezoekers konden op een kaart een datum schrijven, vanaf het jaartal 1851. Langs automatische weg werd dan de titel van het hoofdartikel van de 'New York Times' van de geschreven datum op een kaart afgedrukt.

Om met de leesmachine voor de praktijk bevredigende resultaten te kunnen behalen, moeten bij het schrijven op de te verwerken documenten enige schrijfvoorschriften in acht genomen worden. De fabrikant van de machine beveelt dan ook gebruik in een zgn. 'controlled environment' aan, waarbij de mogelijkheid aanwezig is de deelnemende personen te instrueren. Er zijn enige publikaties over toepassingen van deze machine, waaruit valt af te leiden dat goede resultaten bereikt kunnen worden.

Een andere machine, welke in eerste instantie werd ontwikkeld voor het lezen van machinale afdrukken, zoals schrijfmachineschrift, kan ook met de hand geschreven cijfers lezen. Deze machine werd vervaardigd door de Amerikaanse onderneming Recognition Equipment, welke zich uitsluitend toelegt op het ontwikkelen en produceren van optische leesmachines. De te lezen cijfers behoren met een zacht potlood volgens een model geschreven te worden binnen voorbedrukte vakjes.

De aftastinrichting bestaat uit een matrix van 12 × 48 foto-elektrische cellen. Elk schrijfvakje wordt afgetast met een matrix 12 × 16 fotocellen. Omdat de cellenmatrix van de aftastinrichting groter is dan een schrijfvakje op het document, is een zekere vrijheid toegestaan in de positionering van de schrijfvakjes ten opzichte van de aftastinrichting. Elke fotocel, die 16 verschillende grijstrappen kan onderscheiden, is via een elektrisch netwerk verbonden met 24 omliggende cellen. De waarde van het zwartniveau van al deze omliggende cellen is van belang

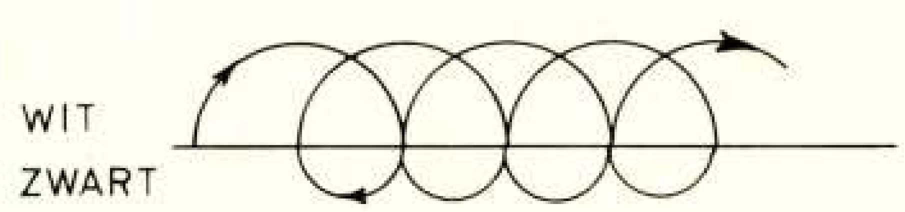


Fig. 4. Detail van de contouraftasting.

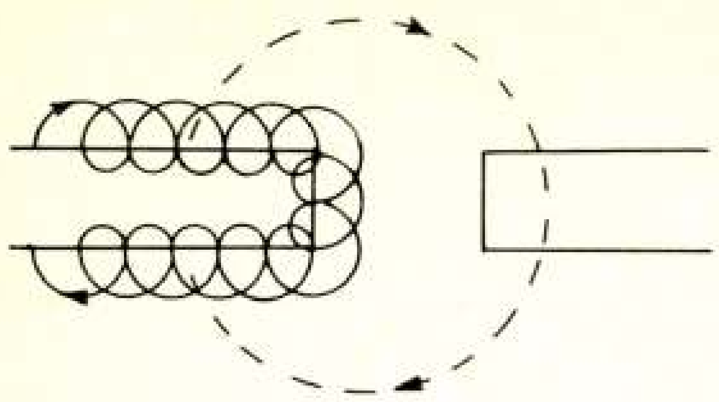


Fig. 5. Het overbruggen van een lijnonderbreking.

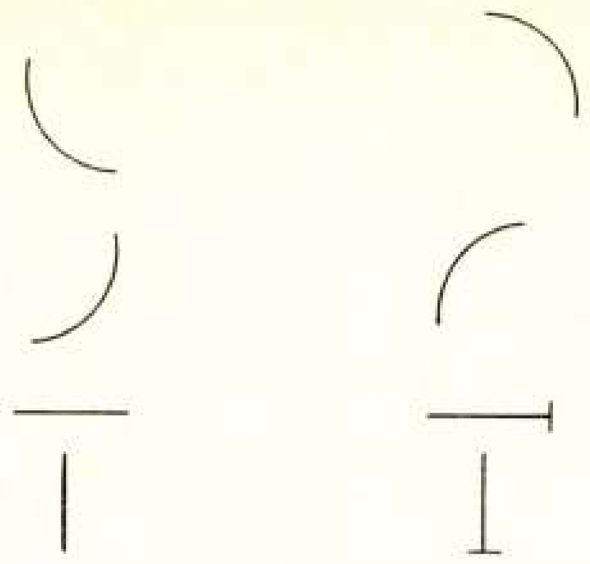


Fig. 7. Karakteristieke kenmerken, waarin geschreven tekens ontleed kunnen worden.

bij de beslissing of een beeldelement als wit of zwart beschouwd moet worden. Hierdoor heeft men bereikt dat de leesresultaten weinig beïnvloed worden door een minder goede afdrukkwaliteit, wanneer gedrukte tekens gelezen worden. Voor de herkenning worden de afmetingen van een cijferafbeelding eerst genormeerd. De matrix van 12 × 16 beeldelementen wordt vervolgens onderverdeeld in 9 horizontale zones, die elkaar overlappen. Per zone vindt een analyse plaats van de waargenomen informatie. Er wordt getest op de aanwezigheid van de karakteristieke kenmerken, die in fig. 7 zijn aangegeven. De onderlinge relatie van deze kenmerken wordt gecodeerd, evenals de positie binnen de matrix. Op grond van deze kenmerken vindt de herkenning plaats. Per seconde kunnen 2400 tekens worden gelezen. Dit geeft een indicatie van de snelheid waarmee de verschillende deelprocessen afgewerkt moeten worden.

3. Recente ontwikkelingen

Bij het toepassen van lezende machines bij administratieve processen heeft men onderkend dat niet alleen de functie van het lezen der informatie van belang is, maar daarnaast ook het gehele verwerkingsproces van de documenten. Dit omvat het papiertransport, een sorteermogelijkheid voor de documenten, de mogelijkheid om de gelezen informatie in de vorm van een code op het document aan te brengen, de mogelijkheid om documenten met niet-leesbare informatie te merken en apart af te leggen, enz. Ook de mogelijkheid, documenten van verschillende afmetingen en van verschillende papierkwaliteiten te verwerken, kan vaak van doorslaggevende betekenis zijn voor de praktische bruikbaarheid van een machine.
De ontwikkeling geeft dan ook als beeld te zien dat men langzamerhand van het denken over een machine overgaat naar het denken in een systeem. Alle bouwstenen van het systeem dienen een grote mate van flexibiliteit te bezitten en ze moeten gemakkelijk kunnen worden aangepast aan de wensen van de gebruikers. Bij de grote, snelle leesmachines vormt het documententransport een niet te verwaarlozen post in het gehele systeem. De verwerkingssnelheid heeft weer vèrgaande consequenties voor de middelen waarmee de aftasting moet plaatsvinden. Ook hangt hiermee de voor de herkenningsprocedure beschikbare tijd samen.

Als een nieuwe ontwikkeling kan genoemd worden het *automatisch lezende eindtoestel.* Recognition Equipment heeft onlangs het optisch lezende eindtoestel, bekend onder de typeaanduiding 'Input 3' op de markt gebracht. Met deze machine, die als tafelmodel is uitgevoerd, kunnen in een langzaam tempo documenten automatisch worden gelezen. Het is mogelijk een aantal van deze lezende perifere toestellen op een centraal opgestelde computer aan te sluiten. Bij de toepassing van dergelijke perifere toestellen onderscheidt men de zgn. IT (*I*ntelligent *T*erminal), waarbij de *herkenning* in het perifere toestel plaatsvindt, en het NIT (*N*on-*I*ntelligent *T*erminal), waarbij in het perifere toestel slechts de *aftasting* van de te lezen informatie plaatsvindt.

De genoemde 'Input 3' is een IT, welke geschikt is voor het lezen van gedrukt en getypt schrift van een aantal gestandaardiseerde schriftsoorten. Ook kunnen gestileerd geschreven cijfers en een paar extra tekens worden gelezen. Om deze mogelijkheid wordt deze machine hier genoemd. In fig. 8 is de mechanische opbouw van de aftastinrichting van de machine getekend. De

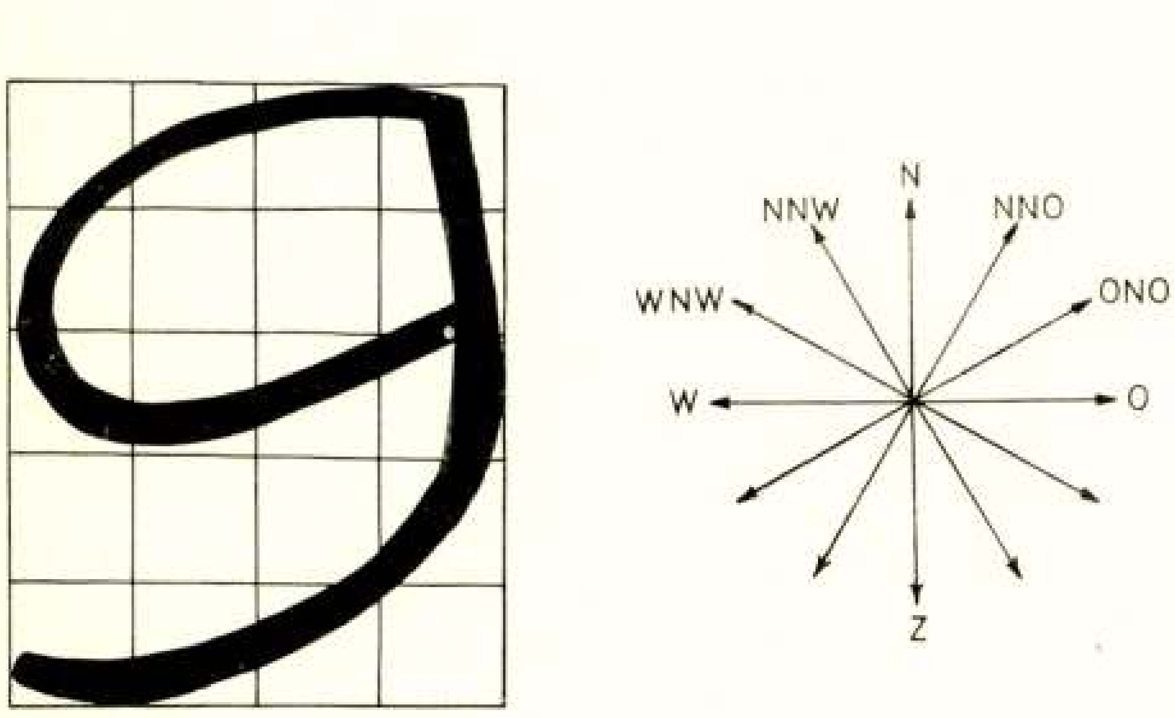


Fig. 6. De normering van een teken en de coderingsmogelijkheden van de hellingen der lijnen.

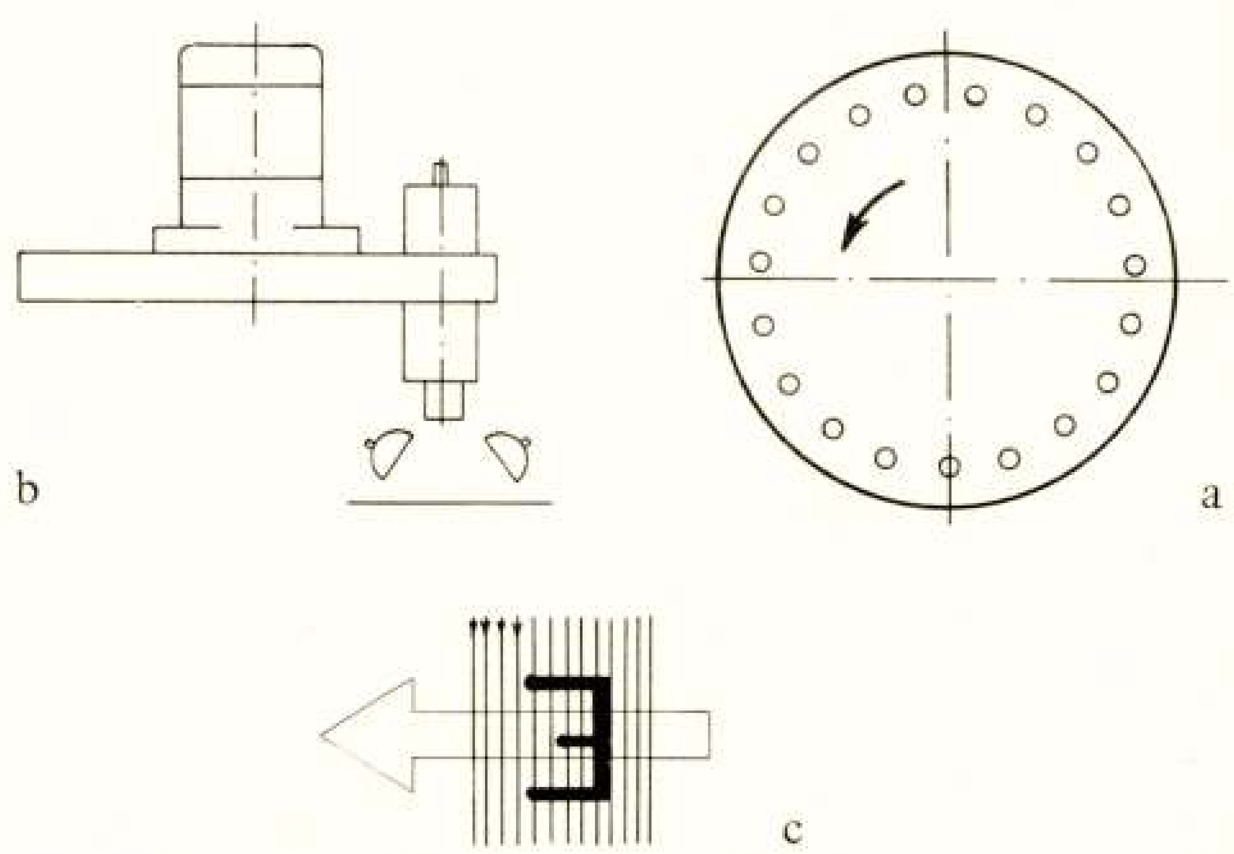


Fig. 8. De aftastinrichting van een optisch lezend eindtoestel.

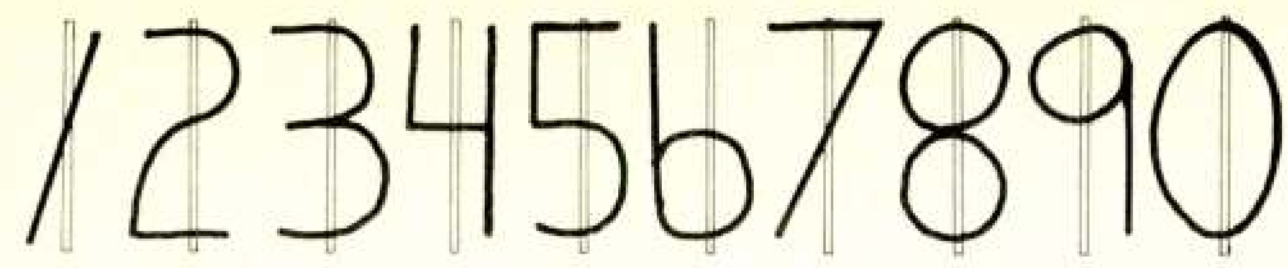


Fig. 9. Gestileerd geschreven cijfers.

aftasting vindt plaats met een door een motor aangedreven schijf met gaatjes (a). Het document wordt verlicht met twee lampen (b). Het gereflecteerde licht dat door een gaatje van de schijf valt, wordt gefocusseerd op een foto-elektrische cel. Wanneer het document in de aangegeven richting wordt getransporteerd (c), en de schijf bovendien een voldoend groot aantal omwentelingen per tijdseenheid maakt, zal elk teken met een aantal, nagenoeg verticale, beeldlijnen worden afgetast. Wanneer de tekens op een regel afgedrukt zijn, gaat de voortbeweging in de langsrichting van die regel. Wanneer een aantal regels op een document afgetast moet worden, wordt na elke regel het document zodanig verschoven, dat dan de volgende regel gelezen kan worden. De gehele aftastprocedure wordt door de processor bestuurd en is dus volledig flexibel. Wanneer documenten met handschrift ingevuld moeten worden, zijn op de daarvoor bestemde plaatsen rode verticale lijnen voorgedrukt. De cijfers moeten nu in zwarte inkt of potlood dóór die lijnen geschreven worden. De cijfers zijn gestileerd, althans zij behoren dat te zijn (fig. 9). De herkenningscriteria zijn zeer eenvoudig: door de verticale streep wordt een teken verdeeld in links en rechts. Voorts kan men een onderste en een bovenste deel aanwijzen. De voornaamste herkenningscriteria zijn:

- het aantal doorsnijdingen in het midden;
- de aanwezigheid van zwarte lijnelementen in elk der kwadranten;
- het wèl of niet gesloten zijn van de lijndelen in elk der kwadranten.

De leessnelheid van deze machine bedraagt ongeveer 40 tekens/s.

4. Een leesinrichting voor niet-gestileerd schrift

Een geheel andere en interessante ontwikkeling heeft zich in Japan voorgedaan. In het kader van de automatisering van de briefpostverwerking heeft het Japanse ministerie voor posterijen kunnen bereiken, dat op alle brieven, briefkaarten, enz., die in Japan worden verstuurd, enige schrijfvakjes op een vaste positie t.o.v. de rand zijn voorgedrukt. In deze schrijfvakjes wordt de 3-cijferige postcode van de plaats van bestemming geschreven in normale, dus 'Arabische' cijfers. De Japanse onderneming Toshiba (afkorting van Tokyo Shibaura Electric Company) heeft een lees-sorteermachine voor poststukken ontwikkeld. Thans zijn reeds in een 20-tal postkantoren in Japan deze lezende sorteermachines in gebruik. Vóórdat men met het ontwerpen van de leesinrichting begon, werd een uitgebreid onderzoek verricht naar de wijze waarop cijfers geschreven worden; 300 000 cijfers werden verzameld, welke door 15 000 personen geschreven waren. Er werd een aftastinrichting gebouwd voor het nauwkeurig aftasten van het te lezen materiaal. De met het lezen verkregen informatie werd voor analyse en het vaststellen van aanwezige kenmerken in een grote computer ingevoerd. Elk schrijfvakje werd hierbij ontleed in $42 \times 60 =$ ruim 2500 beeldpunten; de signaalwaarde van elk beeldpunt werd zodanig gecodeerd, dat 8 verschillende grijsniveaus konden worden onderscheiden. Dit laatste was gewenst, omdat men aan het publiek niet 'speciale' schrijfinstrumenten wilde voorschrijven en omdat men een indruk wilde verkrijgen omtrent de variaties die bij de diverse gebruikelijke inktsoorten in de reflectiecoëfficiënt optreden.

Met behulp van computersimulatie werd een geschikt herkenningssysteem ontwikkeld. Men kan dit in grote lijnen als volgt omschrijven:

Het herkenningsproces is een data-reductieproces waarmee het grote aantal bits dat bij de aftasting is verkregen, voor elk teken gereduceerd wordt tot minder dan 4 bits, om het in te delen in één der tien mogelijke klassen 0 ... 9. Het eerste deel van het toegepaste reductieproces verloopt in een aantal opeenvolgende trappen. Opgemerkt zij dat de reductiefactor per trap 3 à 4 bedraagt, zodat telkens $\frac{1}{3}$ à $\frac{1}{4}$ van de informatie naar een volgende fase wordt doorgegeven. Hierbij doet zich uiteraard het probleem voor wèlke informatie moet worden verwijderd en welke behouden moet blijven om het resultaat minder 'redundant' en dus betrouwbaarder te maken. Het onderzoek heeft geleid tot het realiseren van een lees-sorteermachine voor brieven. Een brief met de daarop voorgedrukte schrijfvakjes wordt met behulp van een televisiecamera afgetast. Het verkregen video-signaal wordt gekwantiseerd, d.w.z. in amplitude begrensd, waardoor slechts de twee waarden *wit* en *zwart* aanwezig blijven. Voorts vindt er een tijdkwantisering plaats. Elk schrijfvakje wordt ontleed in 50×35 beeldelementen. Bij de aftasting worden twee televisiecamera's gebruikt. Bij de ene camera veroorzaken zowel de rode schrijfvakjes als de geschreven cijfers signalen in het video-signaal. Door het toepassen van een optisch filter bij de andere camera bevat het videosignaal alleen de informatie der geschreven cijfers. In het geval dat de cijfers te groot zijn of te dicht bij elkaar zijn geschreven, kan met de aldus verkregen informatie de scheiding of segmentatie der cijfers tot stand worden gebracht. Als de hoogte van een geschreven cijfer een bepaalde waarde te boven gaat, worden de afmetingen van het cijfer gereduceerd. Voorts vindt er een verdunningsproces plaats, waardoor bij dik geschreven cijfers slechts het skelet van het cijferbeeld overblijft. Het resultaat van de hier geschetste voorbewerkingen *normeren* en *verdunnen* is, dat elk cijferpatroon uiteindelijk in een rechthoek van 32×32 beeldelementen past. De informatie wordt in een buffergeheugen geregistreerd en neemt daar 32 woorden of geheugenplaatsen in beslag.

De detectie van karakteristieke eigenschappen verloopt in een aantal deelprocessen. Elk veld van 32×32 beeldelementen wordt in 16 horizontale stroken afgetast met een submatrix van 3×3 bits. Van de 2^9 mogelijkheden binnen de submatrix worden er 7 gebruikt, en wel, om de stand van het lijnstuk in de submatrix in codevorm vast te leggen. Na het aftasten van het gehele beeldveld is voor elk der 16 zones een lijst van 32 waarnemingen aanwezig. In het dan volgende deelproces wordt elke lijst van 32 waarnemingen omgevormd tot een zgn. 'horizontal feature'. Wanneer bijvoorbeeld bij het van links naar rechts opschuiven van de submatrix voor de helling van de lijn achtereenvolgens wordt gevonden: 'positieve helling, aansluitend horizontaal gedeelte, negatieve helling', dan wordt de 'horizontal feature' aangegeven als een bolle kromming. Er worden in totaal 19 verschillende 'horizontal features' gehanteerd; zo kent men bijvoorbeeld ook horizontale en verticale lijndelen, parallel verlopende lijnen, enz. Wanneer de lijst van 16 'horizontal features' is verkregen, volgt het eigenlijke herkenningsproces. Dit bestaat uit het volgtijdelijk vergelijken van de kenmerken van een gelezen cijferpatroon met een aantal

tabellen die in een afzonderlijk kernengeheugen in het geheugen van de machine zijn opgeslagen. Deze tabellen zijn in feite beslissingsnetwerken. Elk netwerk wordt op grond van een gevonden kenmerk in een hogere-orde toestand gebracht. Bij elke toestand is gespecificeerd hoe het netwerk in een volgende toestand kan geraken. Past de reeks van kenmerken van een cijferpatroon bij een bepaalde tabel, dan zal na afloop van het volgtijdelijk vergelijken het bijbehorende beslissingsnetwerk in de eindtoestand geraken, die aangeeft dat het cijferpatroon is herkend. De leesmachine bevat ongeveer 70 diagrammen of netwerken. Met elk diagram kan een klasse der in de praktijk voorkomende cijfervormen herkend worden. Er zijn ongeveer 3500 machinewoorden van elk 32 bits nodig voor het vastleggen van alle diagrammen.

Bij het afwerken van het herkenningsproces wordt de bij de aftasting verkregen lijst van eigenschappen vergeleken met alle toestandsdiagrammen. Wanneer in een toestandsdiagram de volgende eigenschap onjuist is, wordt voor het betreffende toestandsdiagram een strafpunt genoteerd. Als bij het einde blijkt dat meer dan één toestandsdiagram volledig werd doorlopen, wordt als geldig teken dàt gekozen, waarbij het aantal opgelopen strafpunten het geringste is. Omdat volgens deze methode zeer veel wegen doorlopen kunnen worden, kunnen ook zeer veel met de hand geschreven cijfers voorkomende variaties herkend worden. Er zijn ongeveer 10^4 geldige mogelijkheden, hetgeen betekent dat langs gemiddeld 1000 verschillende wegen een eindresultaat bereikt kan worden.

Door de gehele machine uit te voeren als een 'special purpose' computer, waarbij veel bewerkingen in afzonderlijk daarvoor ontworpen elektrische circuits uitgevoerd worden, werd ondanks de gecompliceerdheid van het systeem een redelijk korte verwerkingstijd bereikt. De tijd benodigd voor het herkennen van één cijfer bedraagt 2 à 3 ms.

De praktische resultaten, bereikt bij het op grond van de met de hand geschreven postnummers, sorteren van brieven zijn dat ca. 70% van de toegevoerde brieven uitgesorteerd wordt en dat bij 30% van de brieven één of meer cijfers niet herkend kunnen worden. Verder blijkt dat van de uitgesorteerde brieven ca. 2% in verkeerde sorteervakken terecht komt. Vergeleken bij handbedrijf waar men ook met 1 à 2% missorteringen rekening houdt, is dit niet slecht te noemen.

5. Het onderzoekingswerk in Nederland

In verschillende instituten in Nederland worden onderzoekingen verricht naar de mogelijkheden om met de hand geschreven cijfers automatisch te lezen. Bij het Dr. Neher-Laboratorium van PTT wordt gewerkt aan een leessysteem, waarmee men de verwerking van met de hand geschreven giro-opdrachten hoopt te kunnen automatiseren. Het project kan in een aantal delen worden gesplitst:

a. *Het onderzoek naar een geschikte vorm en indeling van een girokaart*, met name van de schrijfvakjes, zodat deze kaart voor gebruik door een groot publiek acceptabel is. In fig. 10 zijn enige mogelijkheden getekend. Een punt van onderzoek is nog, in hoeverre de vorm en afmetingen van de schrijfvakjes invloed hebben op de duidelijkheid van het schrijven.

b. *Het ontwikkelen van een aftastinrichting*, waarmee de documenten kunnen worden afgetast en waarmee de verkregen informatie naar een processor overgebracht kan worden. Er wordt hierbij gebruik gemaakt van een televisiecamera met een plumbicon-buis. Elk schrijfvakje van een kaart wordt ontleed in 32×32 beeldelementen. De inrichting bevat een drempelschakeling voor het beslissen of een beeldelement wit of zwart is, zodat bij aftasting per schrijfvakje 1024 bits ontstaan.

In de proefopstelling, zoals die thans is gerealiseerd, worden de coördinaten gevormd van die punten waar een contrastovergang optreedt; dit zijn dus de overgangen van wit naar zwart en van zwart naar wit. Wanneer een contrastovergang wordt waargenomen, worden de bijbehorende coördinaten van dat punt in de processor opgenomen. Bij de aftasting ontstaat zodoende een lijst met getallen, die corresponderen met de contouren van een cijferafbeelding. De informatie-overdracht van de gegevens van alle schrijfvakjes op een kaart naar de processor vindt plaats binnen de tijd van één raster, dus binnen 20 ms. Er kunnen in één raster maximaal 16 schrijfvakjes afgetast worden, hetgeen betekent dat ruim 16 000 bits binnen de tijd van 20 ms overgebracht worden naar de processor.

c. *Het bestuderen van herkenningssystemen op in de praktijk verkregen cijfermateriaal*. Vanwege het zeer experimentele stadium waarin alles nog verkeert, zullen de herkenningssystemen uitsluitend programmatisch met behulp van de processor worden beproefd. Er zijn thans enige herkenningsprogramma's in onderzoek, waarbij kenmerken gehanteerd worden zoals samenvloeiingen van lijnen, horizontale en verticale lijndelen, gebogen lijnen, enz. Het zijn deterministische herkenningssystemen, waarin een aantal controlevoorwaarden is ingebouwd.

In het algemeen wordt door onderzoekers op het gebied van de herkenningsproblematiek van met de hand geschreven cijfers de noodzaak gevoeld tot het aanleggen van een verzameling van geschreven cijfers, die dan als referentie zou dienen om de eigenschappen der verschillende herkenningssystemen onderling te kunnen vergelijken. In Amerika bestaan reeds enkele van deze zgn. 'reference sets'.

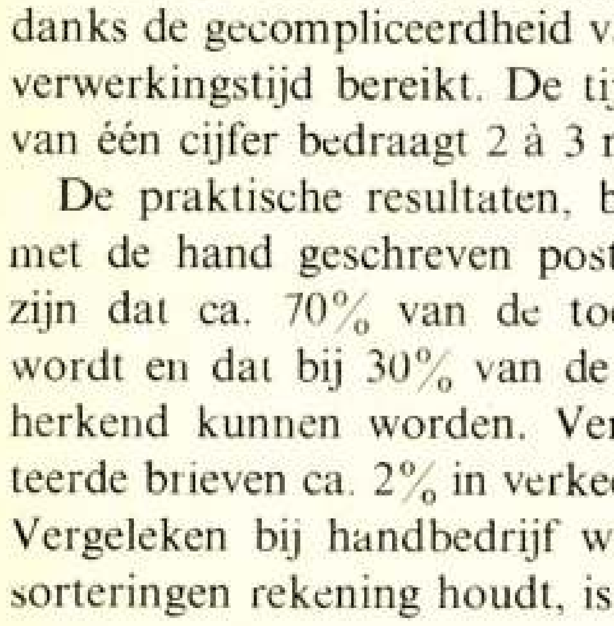


Fig. 10. Voorbeelden van schrijfruimten.

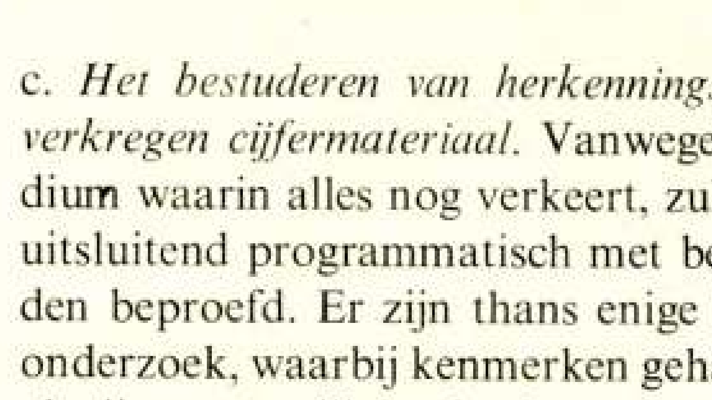


Fig. 11. Enige voorbeelden van voor enkele landen karakteristieke schrijfwijzen.

Er dient opgemerkt te worden dat vele van de karakteristieke kenmerken van het geschreven schrift van land tot land nogal verschillen. Enige bijzonderheden: de wijze waarop men doorgaans in Nederland het cijfer 8 schrijft en zoals het ook op de scholen wordt aangeleerd, is bepaald uniek te noemen. Elders, zelfs in de ons omliggende landen, wordt het cijfer anders gevormd (fig. 11). In deze figuur zijn nog andere voorbeelden opgenomen van karakteristieke schrijfwijzen; zo is aangegeven op welke wijze o.a. in Duitsland veelal het cijfer 1 wordt geschreven. De vorm van het geschreven cijfer 5 volgens fig. 11 treft men vaak in het Japanse handschrift aan.

Om de resultaten die de verschillende onderzoekers bereiken te kunnen vergelijken en om vast te stellen in welke mate de verschillende systemen geschikt zijn voor het herkennen van de Nederlandse schrijfwijze, is het gewenst een afzonderlijke Nederlandse referentieverzameling te vormen. Aan een dergelijke verzameling en aan de wijze waarop deze wordt samengesteld moet een aantal eisen gesteld worden. Wanneer men een enigszins getrouw beeld wil vormen van de oorspronkelijke optische informatie, dan zouden de volgende punten overwogen dienen te worden:

- De keuze van het aantal beeldpunten moet voldoende groot zijn, opdat de door kwantisering veroorzaakte vervorming gering blijft.
- De technische eigenschappen van de aftaster waarmee de informatie verkregen werd, moeten bekend zijn. Hierbij kan men denken aan de spectrale gevoeligheid van de optisch-elektrische omzetter zowel als aan de spectrale energieverdeling van de lichtbron. Voorts zou het gewenst zijn als er een lineair verband zou bestaan tussen de intensiteit van het op de aftaster vallende licht en het elektrische signaal, dat door de aftaster wordt afgegeven.
- De waarde van de reflectiecoëfficiënt van elk beeldpunt, de grijstrap, moet eveneens in gecodeerde waarde opgegeven worden. Dit laatste is gewenst om elk teken als een driedimensionale figuur te kunnen beschouwen, waarbij in de Z-richting de zwartingsgraad aangegeven wordt. Wanneer slechts binaire informatie verstrekt zou worden, is één doorsnede van het teken gemaakt en zou men een resultaat verkrijgen dat afhankelijk is van de werking van een zwart-wit discriminatorschakeling.
- De referentieverzameling dient representatief te zijn. Men mag aannemen dat giro-opdrachten en bankoverschrijvingen in het algemeen vrij duidelijk en zorgvuldig geschreven zijn. De numerieke gegevens die op dit soort documenten geschreven zijn zullen wellicht een ander beeld geven dan geheel willekeurig geschreven cijfers.

Tenslotte zij opgemerkt dat het in een later stadium wellicht nuttig kan blijken om verschillende referentieverzamelingen te vormen, welke elk afzonderlijk representatief zijn voor een categorie van personen die onder bepaalde omstandigheden hun gegevens verstrekken.

Literatuur

[1] SHEINBERG, I.: The 'Input 2' document reader. (A new optical character recognition system.) Pattern Recognition. Vol. 2, No. 3, September 1970.

[2] HOLT, A. W.: Smart terminals. Datamation, October 1970.

[3] GENCHI, H., MORI, K., WATANABE, S. and KATSURAGI, S.: Recognition of handwritten numeral characters for automatic letter sorting. Proceedings of the IEEE. Vol. 56, No. 8, August 1968.

[4] MORI, K., GENCHI, H., WATANABE, S. and KATSURAGI, S.: Microprogram controlled pattern processing in a handwritten mail reader-sorter. Pattern Recognition. Vol. 2, No. 3, September 1970.

IV. De ontwikkeling van een optische lees- en sorteermachine

door ir. H. van Steenis en ir. A. Kempkes, IBM-Laboratorium, Uithoorn

Synopsis: *Development of an Optical Reader Sorter.*
The IBM 1275 Optical Reader Sorter was designed to read numerals and symbols of either the OCR-A or the OCR-B type font, allowing for wide print quality tolerances.

At first a brief discussion is given in this article of the environment for the reader and the consequence for its design. Then the concept and philosophy for the two-dimensional integrating recognition logic are treated more in detail. Finally a short description is given of the computer simulation and the techniques used.

1. Inleiding

De snelle toeneming van het aantal financiële transacties, veroorzaakt door een in voortdurende groei verkerende economie, heeft de banken en andere financiële instellingen ertoe gedwongen uit te zien naar een wijze waarop de behandeling van cheques en overschrijvingen kan worden geautomatiseerd. Voor dit doel is ongeveer tien jaar geleden een leessysteem ontwikkeld voor magnetisch schrift (MICR = *M*agnetic *I*nk *C*haracter *R*ecognition).

In de Verenigde Staten werd het *E*-13B schrift ontwikkeld en in Europa het *CMC*-7 schrift. Beide schriftsoorten eisen speciale kostbare drukapparatuur en een hoge mate van controle op de drukkwaliteit.

In Europa – waar in tegenstelling tot de Verenigde Staten het lezen van magnetisch schrift weinig wordt toegepast – is een sterke neiging merkbaar om over te gaan tot het optisch lezen van cheques en overschrijvingen (OCR = *O*ptical *C*haracter *R*ecognition). Het hoofdmotief hiervoor is dat snelle regeldrukkers kunnen worden gebruikt voor het prepareren van de formulieren zodat kostbare drukapparatuur, noodzakelijk bij magnetisch lezen, kan worden vermeden. De IBM-1275 optische lees- en sorteermachine werd tegen deze achtergrond ontwikkeld.

2. Eisen

De volgende eisen werden gesteld vanuit het toepassingsgebied van de optische lezer.

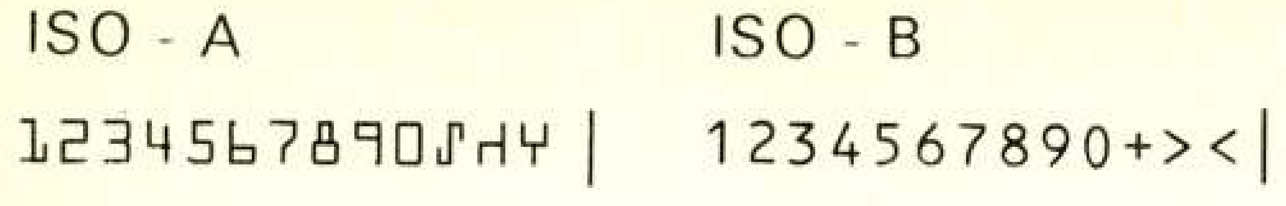


Fig. 1. Afbeelding van de beide beschikbare schriftsoorten: ISO-A en ISO-B.

a. *Schriftsoort*

De optische lezer moet in staat zijn een algemeen aanvaarde schriftsoort te lezen, zodat formulieren tussen de diverse financiële instellingen kunnen worden uitgewisseld. De keuze tussen ISO-A en ISO-B schrift (fig. 1) is aan de gebruiker overgelaten (ISO = *I*nternational *S*tandardizing *O*rganization). Herkenningslogica is voor beide schriftsoorten ontwikkeld.

b. *Leessnelheid*

Een hoge leessnelheid is een vereiste. Een zeer groot aantal formulieren moet per dag kunnen worden gelezen. Bovendien moeten de formulieren worden gesorteerd, wat meer dan één sorteergang door de machine met zich meebrengt.

In het algemeen wordt de te lezen informatie in één regel samengebracht, die in één gang door de optische lezer kan worden gelezen. De transportsnelheid van de formulieren bedraagt 6,7 m/s.

c. *Leesprestaties*

De doorvoersnelheid van de optische lezer is maximaal 96 000 formulieren van vijftien centimeter lengte per uur. In het geval dat van elk formulier dertig tekens (cijfers en speciale symbolen) moeten worden gelezen, komt dit neer op ongeveer drie miljoen gelezen tekens per uur, ofwel ~ 800 tekens per seconde. De leesprestaties van de lezer moeten dus zeer goed zijn, om het aantal leesfouten per tijdseenheid te beperken.

d. *Drukkwaliteit*

Een grote variëteit aan drukmachines en het feit dat de formulieren ontstaan in verschillende locaties, maakt het waarschijnlijk dat er weinig controle kan zijn op de drukkwaliteit. Om die reden moest er een nauwkeurig herkenningssysteem worden ontworpen, zodat ook druk van minder goede kwaliteit kan worden gelezen. Met deze eisen als uitgangspunt werd het IBM-1275 herkenningssysteem [1] ontwikkeld en gekoppeld aan het transportsysteem van de IBM-1419 magnetische lees- en sorteermachine. De volledige IBM-1275 is afgebeeld in fig. 2.

Het compartiment aan de linkerzijde bevat de herkenningslogica. Formulieren worden aan de rechterzijde in de machine gebracht. Het optisch aftasten vindt plaats in het ernaast gelegen deel. In het middengedeelte van de machine zijn twaalf aflegvakken zichtbaar, die worden gebruikt tijdens het lezen of sorteren, een dertiende aflegvak is voor de afgewezen formulieren. Door middel van schakelaars boven de aflegvakken kan de operator de gewenste functie selecteren.

3. Lees- en herkenningssysteem

In het lees- en herkenningssysteem van IBM-1275 kan men zes basiselementen onderkennen (fig. 3). De leeskop zet variaties in het gereflecteerde licht om in elektrische signalen. De leesversterkers versterken deze signalen. De versterkte signalen worden hierna via een drempelcircuit gedigitaliseerd in zwarte en witte rasterpunten. In de matrix worden de rasterpunten verzameld ten einde een elektronisch beeld op te bouwen van dat gedeelte van het formulier, dat net is afgetast.

De vergelijkingslogica analyseert voortdurend het elektronische beeld en bepaalt op het juiste tijdstip de mate van gelijkenis tussen dit elektronische beeld van een teken en elk van de tekenklassen die kunnen worden herkend. De beslissingslogica ten slotte inspecteert de uitkomst van de vergelijkingen en besluit of het teken moet worden afgewezen dan wel is herkend. In het laatste geval wordt het teken geïdentificeerd.

Elk van deze basiselementen zal nu meer in detail worden behandeld.

3.1. *De leeskop*

Van groot belang voor de herkenningsprestaties is de resolutie (het oplossend vermogen) van de leeskop. De resolutie bepaalt

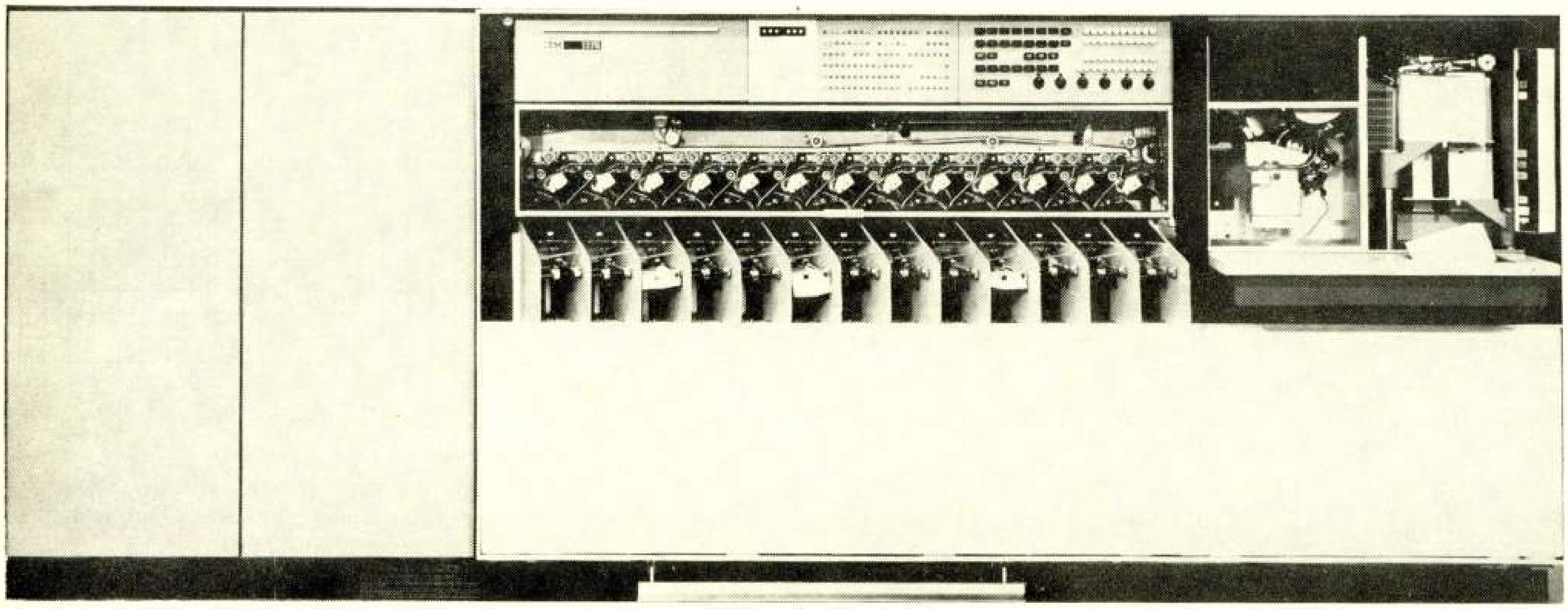


Fig. 2. De IBM-1275 optische lees- en sorteermachine.

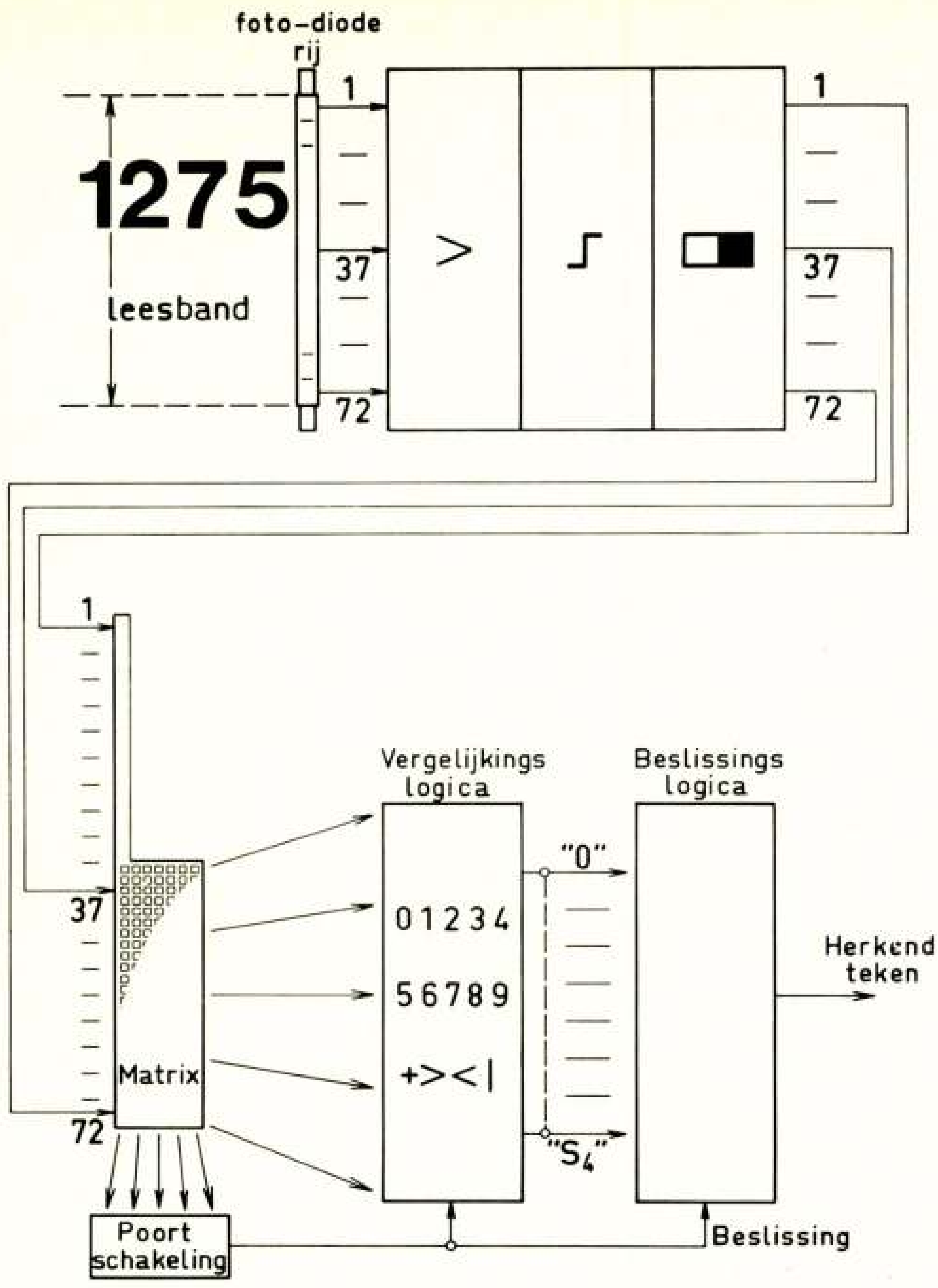


Fig. 3. Schema van het IBM-1275 lees- en herkenningssysteem.

het kleinste detail dat nog kan worden waargenomen. Op grond van ervaringen met voorgaande leessystemen is gekozen voor een resolutie van 0,127 mm. Een elementaire cel in de matrix representeert derhalve een gebiedje van 0,127 bij 0,127 mm. De samenstelling van de optische aftaster is conventioneel:



Fig. 4. Gedrukte tekens en de resulterende matrixpatronen (de duur van de vergelijkingsperiode is weergegeven onder de figuren).

twee wolframlampen verlichten dat gedeelte van het formulier, dat zich juist onder de leeskop bevindt; met behulp van een lens wordt op een rij van 82 fotodiodes een 1 : 1 afbeelding gevormd van het af te tasten gebied.

De steek van twee aangrenzende fotodiodes is gelijk aan de eerder genoemde 0,127 mm. Van deze rij fotodiodes vormen de binnenste 72 fotodiodes het gebied dat het te lezen teken moet passeren, de zgn. leesband. Deze band is ongeveer driemaal de hoogte van een teken ten einde praktische druk- en snijtoleranties te kunnen toelaten.

De tien resterende fotodiodes, vijf aan de onderzijde en vijf aan de bovenzijde, worden gebruikt om onregelmatigheden zoals annuleringsstempels of scheef gedrukte regels te detecteren.

Een voordeel van de combinatie silicium fotocel/wolframlamp is dat de spectrale gevoeligheid der cel in het nabije infrarood ligt, zodat de leeskop blind is voor lichtgekleurde achtergrondbedrukking, maar wel gevoelig blijft voor bepaalde inkten, toegepast in de gebruikte linten. Zie als voorbeeld fig. 4a, waarin de afdruk van een bruin stempel is weergegeven.

3.2. *De leesversterkers*

Elke fotodiode is verbonden met een eigen logaritmische versterker. De belangrijkste eigenschap van een logaritmische versterker is dat de verandering van de uitgangsspanning uitsluitend afhankelijk is van de afgetaste contrastverandering en onafhankelijk van de verlichtingssterkte en de gevoeligheid van de fotodiodes. Door het 'witniveau' van iedere leesversterker op een bepaalde waarde vast te leggen kan een zeer goede uniformiteit in signaal van de leeskanalen onderling worden bereikt.

3.3. *Het drempelcircuit*

Het signaal van iedere versterker wordt vergeleken met een dynamische drempelspanning en wordt afhankelijk van de polariteit van het resultaat als zwart of wit beoordeeld. De dynamische drempelspanning moet op een correcte wijze worden bepaald ten einde een zo natuurgetrouw mogelijk beeld van het gedrukte teken te verkrijgen.

Er zijn drie factoren, die de drempelspanning van een leeskanaal beïnvloeden. De drempelspanning kan niet beneden een bepaald niveau dalen (de minimumdrempel); dit om te voorkomen dat onregelmatigheden in het papier worden geïnterpreteerd als een zwart/wit-patroon. De drempelspanning wordt beïnvloed door het gedetecteerde contrast in nabij gelegen leeskanalen. De beïnvloeding is zwakker naarmate de afstand tot het leeskanaal groter is. Tenslotte beïnvloedt het sterkste zwartsignaal, gedetecteerd in de rij fotodiodes, de drempelspanning. De gewenste mate van beïnvloeding is experimenteel bepaald.

De resultaten, bereikt met dit systeem, zijn weergegeven in fig. 4 (b, c, d en e). De afgedrukte tekens in fig. 4c zijn moeilijk goed te reproduceren, de zwarting van de tekens is juist onder de gestelde grens.

De uitgangsspanning van elk leeskanaal – welke binair is in amplitude en continu in tijd – wordt elke 19 µs bemonsterd. Dit komt neer op een bemonstering van elke 0,127 mm op het formulier.

Het resultaat van een bemonstering is een kolom van 72 binaire rasterpunten. De informatie in de kolom wordt bit voor bit naar de matrix gestuurd. Elke bit stelt een gebiedje voor van 0,127 bij 0,127 mm op het papier; door strooiing en dergelijke zal het werkelijke gebiedje iets groter zijn.

3.4. *De matrix*

Bits afkomstig van de leeskanalen worden continu verzameld in de matrix, waar een elektronisch beeld wordt gevormd van de 2,5 mm die zojuist is afgetast. De matrix is een 720 bits lang schuifregister (klokfrequentie ongeveer 1 900 000 Hz), dat het te herkennen patroon in alle mogelijke posities brengt.

Het 720-bits schuifregister kan worden voorgesteld door 20 kolommen, corresponderend met evenzoveel bemonsteringen over 2,54 mm met een hoogte van 36 bits, de helft van het aantal fotodiodes. Omdat een teken slechts eenderde van de hoogte van de fotodiode-rij kan beslaan, wordt het bitpatroon van de bovenste 36 fotodiodes gesuperponeerd op het bitpatroon van de onderste 36 fotodiodes.

3.5. *De vergelijkingslogica*

Verbonden met de matrix is een aantal elektronische maskers – voor elke te herkennen teken-klasse ongeveer vijf – die kunnen worden vergeleken met het bitpatroon in de matrix, afkomstig van het juist afgetaste teken. Deze vergelijking dient slechts dan plaats te vinden wanneer het bitpatroon ongeveer goed gepositioneerd is ten opzichte van de maskers.

Een speciale poortschakeling maakt alleen dan vergelijking mogelijk wanneer de waarschijnlijkheid van een goede positionering maximaal is. De goede werking van deze poortschakeling is van het hoogste belang. De schakeling mag niet reageren op kleine vlekken afkomstig van stof of vuil op het formulier, maar moet wel in alle gevallen aanspreken op smalle tekens met eventueel missende delen. Het feit dat de breedte van de tekens varieert, dat een deel van de tekens slechts gedeeltelijk is afgedrukt en dat de afstand tussen de tekens onderling sterk kan variëren (fig. 4f) maakt de poortschakeling complex; op details van deze schakeling zal hier niet worden ingegaan.

Het begin van de vergelijkingsperiode wordt bepaald door de afstand tot het voorgaande teken en door de minimumvoorwaarden, gesteld aan een patroon. De duur van de vergelijkingsperiode wordt bepaald door de breedte van het teken en de afstand tot het volgende teken. Afhankelijk van de breedte van het teken verplaatst het bitpatroon zich in de vergelijkingsperiode over 6 ... 13 kolommen. Ter vergelijking: de nominale breedte van een teken is 14 kolommen, de nominale breedte van de witte ruimte tussen 2 tekens is 6 kolommen.

Gedurende de vergelijkingsperiode, dit is gemiddeld gedurende $9 \times 36 = 324$ klokpulsen, wordt het bitpatroon vergeleken met de elektronische maskers. Het aantal malen dat tenminste één van de maskers aanspreekt op het bitpatroon wordt geteld, om aan het einde van de vergelijkingsperiode een beslissing mogelijk te maken over de identiteit van het afgetaste

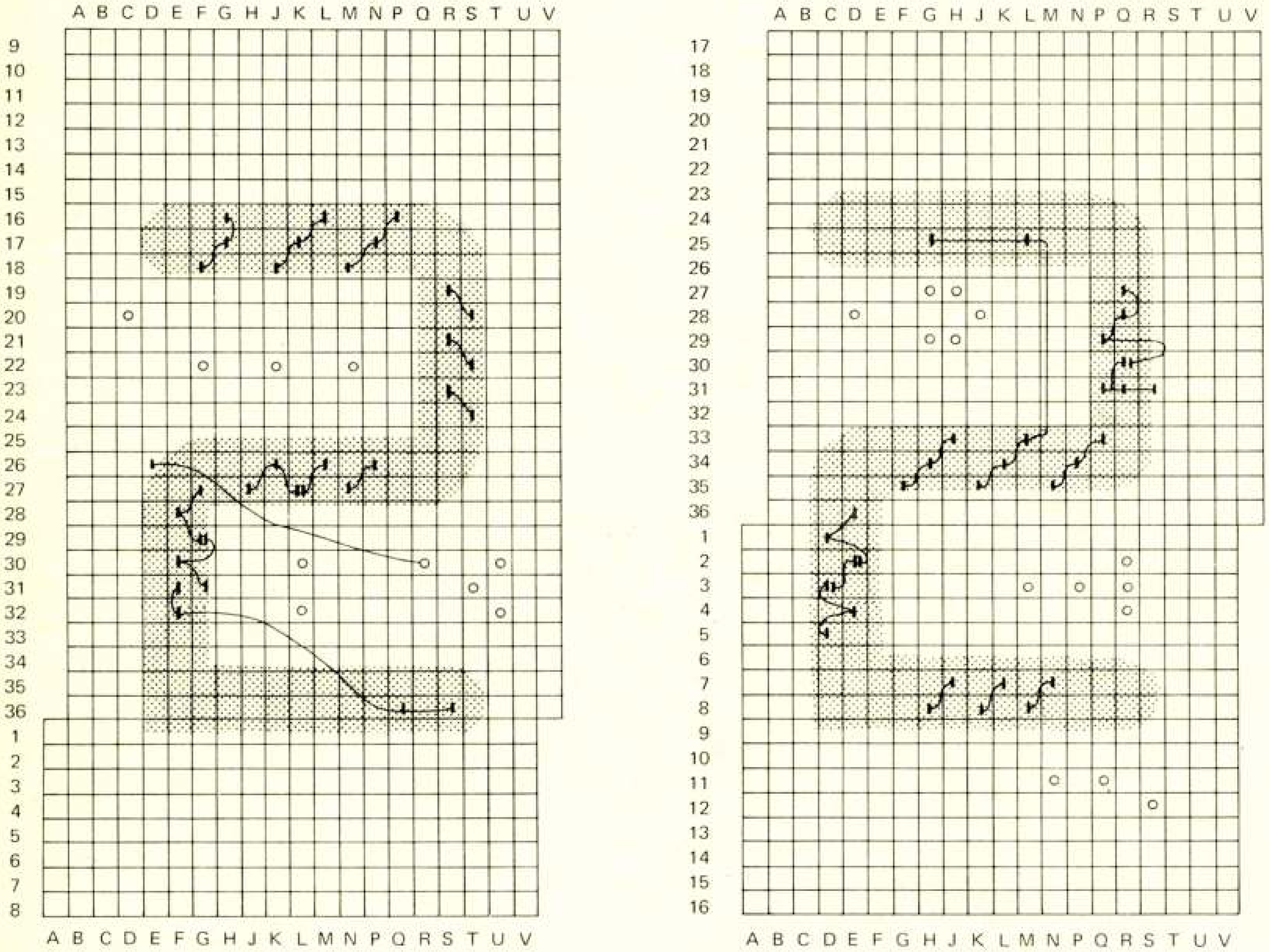


Fig. 5. Twee mogelijke maskers voor een '2' in het ISO-A schrift. Links: masker voor een '2', waarvan de onderzijde kan ontbreken. Rechts: masker voor een '2', waarvan de bovenzijde voor een groot deel kan ontbreken.
Legende: ▌ = matrixpositie moet zwart zijn
○ = matrixpositie moet wit zijn
⊢⊣ = tenminste één van de door de lijn verbonden matrixposities moet aan de aangegeven conditie voldoen

teken. Het doel van elk masker is zo vaak als mogelijk aan te spreken op de juiste patronen. Overeenstemming tussen het masker en patronen van andere tekenklassen dient zoveel mogelijk te worden vermeden.

Tijdens de ontwikkeling van de maskers is het aantrekkelijk gebleken voor elke teken-klasse een aantal alternatieve maskers te gebruiken. In het algemeen zijn er 5 maskers voor elke te herkennen teken-klasse ontworpen, nl. voor patronen zonder linkerzijde, zonder rechterzijde, zonder bovenzijde en voor patronen met een grote streepdikte. In fig. 5 zijn 2 mogelijke maskers voor de '2' van het ISO-A schrift weergegeven. Gebleken is dat de op op een dergelijke wijze vastgelegde maskers voldoende vrijheid bieden om problemen op te lossen.

3.6. *Beslissingslogica*

Tijdens de vergelijkingsperiode wordt het aantal malen, dat tenminste één van de maskers aanspreekt op het bitpatroon, geteld; er is één teller per teken-klasse (dus in totaal zijn er 14 tellers). Aan het einde van de vergelijkingsperiode worden de tellers geïnspecteerd en wordt de beslissing genomen het teken te identificeren dan wel af te wijzen.

De tellingen moeten aan twee criteria voldoen wil het juist afgetaste teken positief geïdentificeerd worden:

1. De hoogste telling moet een zekere waarde overschrijden; in ons geval is deze waarde 4.

2. Het verschil tussen de hoogste telling en de op één na hoogste telling moet ook een zekere waarde overschrijden, in ons geval eveneens 4. In principe kunnen deze waarden verschillend zijn.

Als aan beide criteria is voldaan, geeft die teller, die de hoogste telling bevat, de teken-klasse aan die herkend is. Het patroon wordt afgewezen als aan één van beide criteria niet is voldaan. In het laatste geval kan nog onderscheid worden gemaakt tussen een conflict (1, maar niet 2) en een zuivere afwijzing (niet 1). Tellingen (tussen haakjes) en beslissingen op een aantal patronen zijn weergegeven in fig. 6.

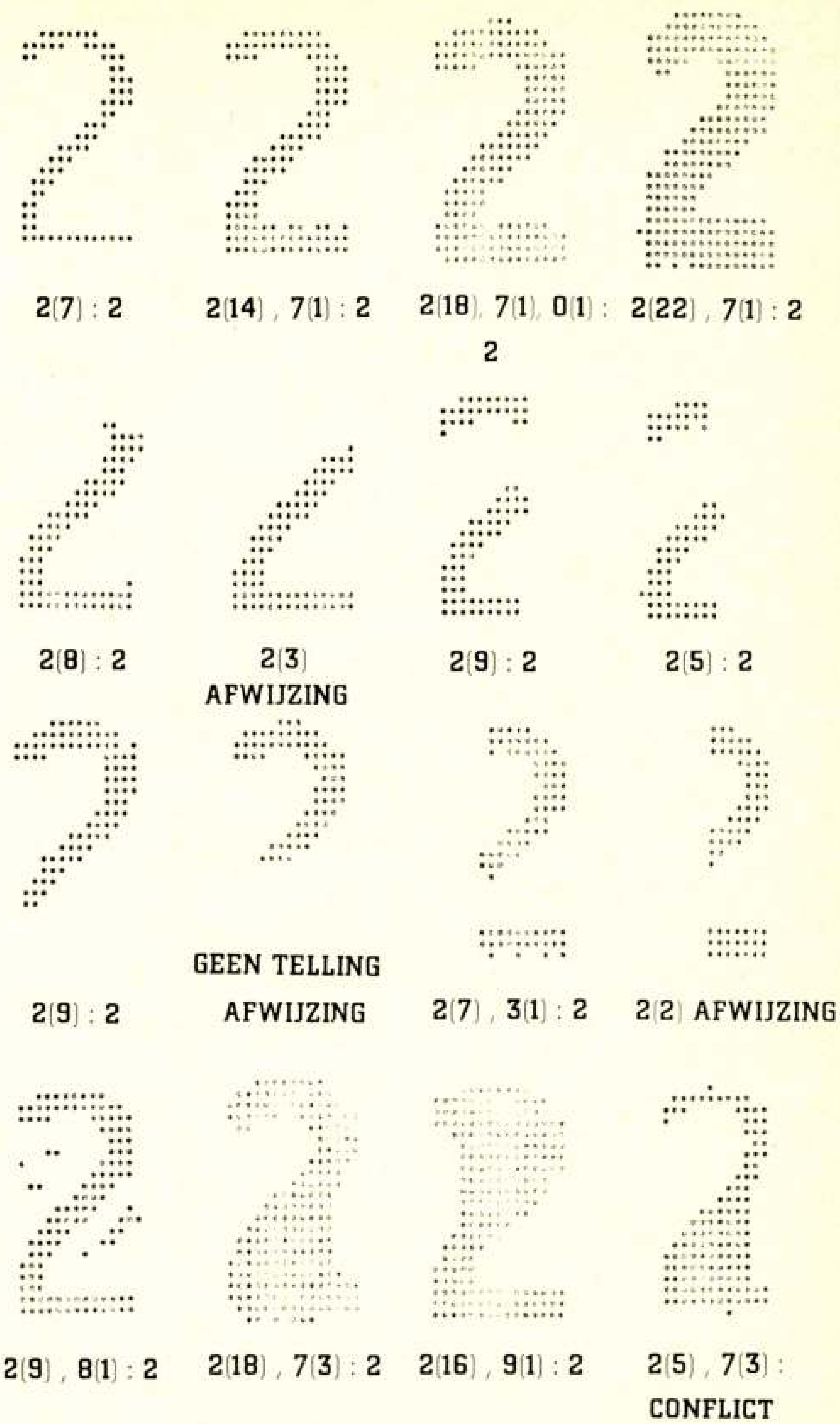


Fig. 6. Verschillende patronen van het teken '2' (ISO-B schrift) samen met de resulterende tellingen en beslissingen.

3.7. *Het herkenningssysteem*

Ten einde de eigenschappen van het besproken herkenningssysteem beter te kunnen waarderen, is een vergelijking met het conventionele matrix-herkenningssysteem nuttig. In deze conventionele systemen (bijv. de IBM-1418 en de IBM-1428 lees- en sorteermachines [2, 4]) is een eenmalige overeenstemming tussen het bitpatroon en een masker voldoende voor een herkenning. Indien twee maskers aanspreken op hetzelfde patroon ontstaat een conflict en het patroon moet worden afgewezen.

Een eerste stap tot de verbetering van het conventionele systeem is het toevoegen van tellers, waarmee het aantal overeenstemmingen tussen het patroon en de verschillende maskers kan worden geteld. Omdat de elektronische maskers niet zijn gewijzigd, zal in het beste geval een aantal van de conflicten kunnen worden opgelost. Inderdaad bleek, dat op deze manier 80% van de op de IBM-1418 optredende conflicten kon worden opgelost zonder dat er herkenningsfouten ontstonden.

De tweede stap ter verbetering is een andere doelstelling bij het ontwerpen van de maskers. In plaats van exclusieve maskers, die uitsluitend aanspreken op de juiste patronen, moeten nu maskers worden ontwikkeld, die zo vaak mogelijk aanspreken op de juiste patronen en zo min mogelijk op patronen van andere teken-klassen. Dit betekent een aanzienlijke vereenvoudiging van dat deel van de maskers, dat moet voorkomen dat onjuiste patronen aanspreken.

Een bijkomend voordeel van de beslissing op grond van tellingen is dat de verhouding tussen afwijzingen en substituties kan worden gevarieerd zonder dat de elektronische maskers behoeven te worden veranderd. Indien de beide telcriteria worden verhoogd, neemt het aantal substituties af en het aantal afwijzingen toe. Verlaging van de telcriteria veroorzaakt het omgekeerde. Het optimale telcriterium voor de 1275 werd vastgesteld met het oog op de verhouding tussen de kosten van het corrigeren van een substitutie en een afwijzing.

De met het besproken herkenningssysteem behaalde resultaten zijn zeer goed. Het is onmogelijk om deze resultaten in getallen weer te geven zonder daarbij zeer nauwkeurig de drukkwaliteit en de kwaliteit van de formulieren te specificeren. Onder actuele omstandigheden is de drukkwaliteit niet meer alleen de door-

slaggevende factor voor de leesprestaties van het systeem. Met andere woorden: papierdefecten, stempels en incorrecte positionering van de bedrukking op het formulier kunnen in hoge mate de leesprestaties beïnvloeden.

4. Het instrumentatiesysteem

De ontwikkeling van de herkenningslogica zou onmogelijk zijn geweest zonder een goed instrumentatiesysteem; dit geldt te meer wanneer de drukkwaliteit op de te lezen formulieren zo sterk kan variëren als dit het geval kan zijn bij de 1275. Tijdens de ontwikkeling is een instrumentatiesysteem toegepast dat analoog is aan dat wat wordt gebruikt voor de IBM-1975 paginalezer, ontwikkeld voor de administratie van de sociale verzekering in de Verenigde Staten [3].

Het uiteindelijke ontwerp voor de elektronische maskers werd bereikt via een iteratief optimaliseringsproces onder controle van de ontwerper. Zijn hulpmiddelen daarbij waren een computer, een complete IBM-1275 lezer en een verzameling programma's.

Tijdens de beginfase van de ontwikkeling is het aantal tekens, dat moet worden getest, nog gering. Zodra de maskers verfijnder en intelligenter worden, groeit dit aantal al snel tot enkele miljoenen. De efficiency en de nauwkeurigheid van het instrumentatiesysteem zijn dan van vitaal belang.

Een computer – een Systeem/360 Model 40 met een 256k geheugen – is het hart van het instrumentatiesysteem. Hiermee verbonden is een complete 1275 optische lezer, voorzien van een speciale aanpassingseenheid, die de computer rechtstreeks verbindt met verschillende delen van de herkenningslogica. Via deze eenheid kunnen matrixpatronen samen met gegevens over de herkenning – zoals de tellingen – naar de computer worden gezonden ten einde daar op magneetbanden te worden vastgelegd. Naast dit instrumentarium is een aantal programma's ontwikkeld, bijv. om statistische gegevens te verkrijgen, om matrixpatronen te rangschikken en om herkenningslogica te controleren. Een herkennings-simulatieprogramma, waarin de herkenning zich exact gedraagt als in de IBM-1275 lezer, is een ander essentieel element van het instrumentatiesysteem. De elektronische maskers worden vertaald in Booleaanse logica en worden verponst aan het simulatieprogramma aangeboden.

Het begin van een nieuwe iteratie is in alle gevallen: het lezen van een groot aantal formulieren door de 1275 (fig. 7). De 1275 is dan voorzien van maskers volgens het laatste iteratieniveau. De formulieren moeten zoveel mogelijk een representatieve afspiegeling vormen van de variëteit aan formulieren, die onder normale omstandigheden optreedt. De formulieren worden gelezen, statistische gegevens worden verzameld en de patronen van niet-herkenbare tekens worden op magneetband 'geschreven', gerangschikt en afgedrukt. De ontwerper van de maskers bestudeert de afgedrukte patronen en tellingen en tracht de maskers te verbeteren. Modificaties in de maskers worden verponst en met het herkennings-simulatieprogramma gesimuleerd, eerst op de zojuist verzamelde patronen, later ook op probleempatronen die eerder zijn verzameld.

De maskers worden gemodificeerd en gesimuleerd totdat uiteindelijk een bevredigend masker is ontworpen. Deze maskers worden in de 1275 aangebracht. Met behulp van de ver-

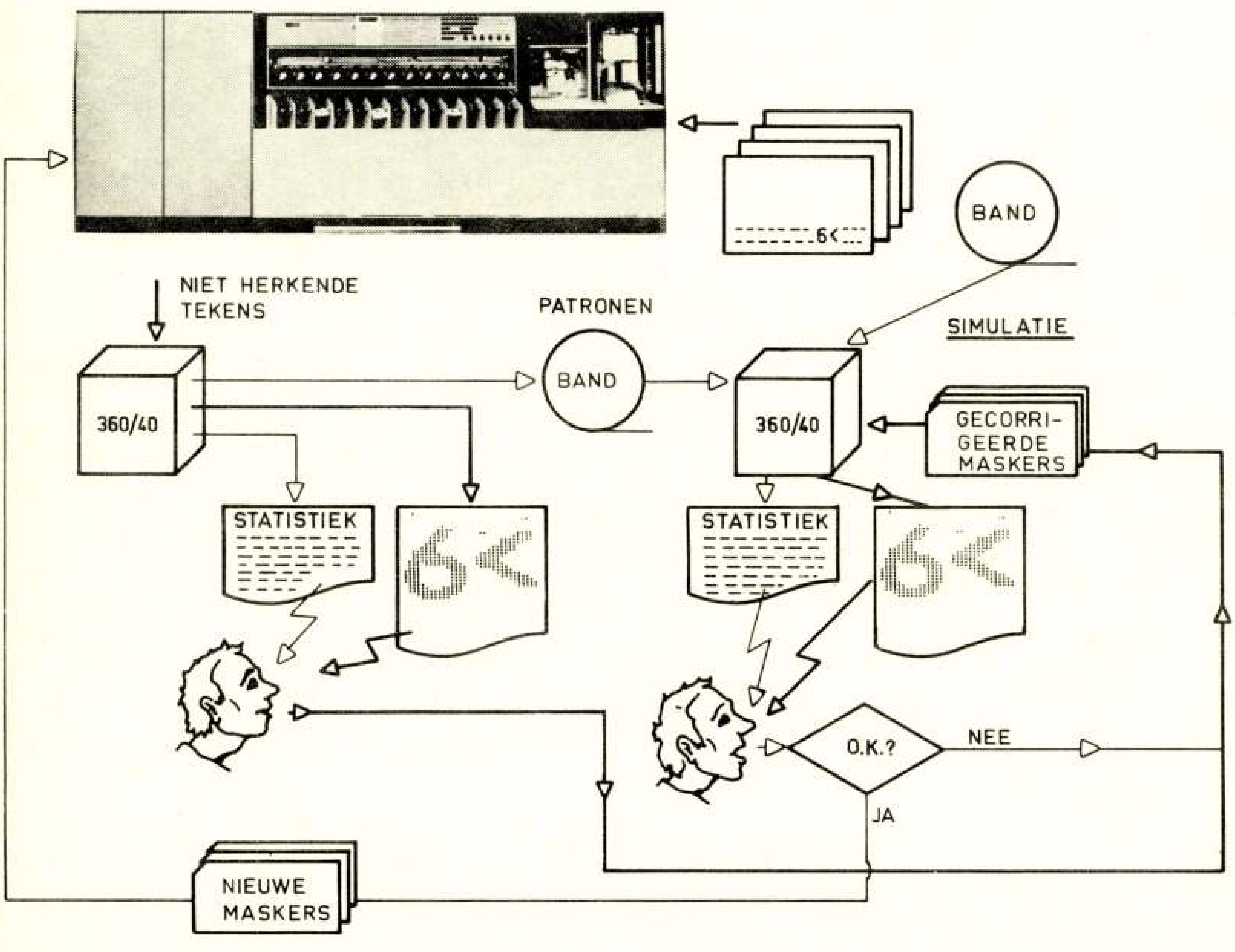


Fig. 7. Het instrumentatiesysteem, gebruikt bij de ontwikkeling van de herkenningslogica.

ponste maskers, gebruikt bij de laatste simulatie, wordt een groot aantal testpatronen ontworpen, die kunnen worden gebruikt bij het foutenvrij maken van de zojuist in de 1275 aangebrachte maskers. Deze procedure versnelt het vinden en herstellen van fouten enorm en verzekert dat de in de 1275 aangebrachte maskers identiek zijn aan de tijdens de laatste simulatie gebruikte maskers. Zo nodig kan dan met een nieuwe iteratie worden begonnen.

5. Slot

Uitgaande van het programma van eisen, gesteld door het toepassingsgebied, is beschreven hoe het IBM-1275 herkenningsprogramma is ontwikkeld. De nadruk is gelegd op het feit dat het leessysteem in staat moet zijn ook zeer slecht afgedrukte tekens te behandelen.

Tot het goed functioneren van het leessysteem hebben de volgende factoren het meest bijgedragen:

- Het zelfregelende leessysteem en de hoge resolutie in de matrix, beide noodzakelijk om ook zeer slechte druk nog correct te kunnen reproduceren.
- Het herkenningssysteem, dat geen exacte registratie van het patroon in de matrix vereist en waarin de herkenning is gebaseerd op de mate van overeenkomst tussen het patroon en de diverse elektronische maskers.
- Het instrumentatiesysteem, dat is gebruikt tijdens de ontwikkeling van de herkenningslogica.

De '1275' optische lees- en sorteermachine werd ontwikkeld in samenwerking met de IBM-laboratoria in Rochester (Minnesota) en Endicott (New York).

6. Literatuur

[1] 'IBM 1275 Optical Reader Sorter'. IBM Systems Reference Library A 19-0034.

[2] LEIMER, J. J.: Design factors in the development of an optical character recognition machine. IRE Transaction on Information Theory, Vol. I T-8, no. 2, pp. 167 ... 171 (February 1962).

[3] IBM Journal of Research and Development, Vol. 12, no. 5, pp. 346 ... 372 (September 1968).

[4] FONTIJNE, P.: Principe van IBM machines voor optische tekenherkenning. 'De Ingenieur', **82**, nr. 43 (23 oktober 1970).

Korte technische berichten

WARC-ST, 1971

Van 7 juni ... 17 juli 1971 werd in Genève een Wereld Administratieve Radio Conferentie gehouden, welke beperkt was tot de satelliettelecommunicatie. Deze conferentie, belegd door de Internationale Telecommunicatie Unie, en gemakshalve aangeduid als WARC-ST, werd nodig geacht om bij de snelle ontwikkeling van de telecommunicatie met en via satellieten de voorzieningen, getroffen in het radioreglement in 1963 betreffende het gebruik van frequenties voor satellietcommunicatie, aan te passen en uit te breiden.

De uitbreidingen hebben vnl. betrekking op het gebruik van het frequentiespectrum boven 10 GHz. In het gebied van 40 GHz ... 275 GHz werd een eerste verdeling gemaakt voor uitsluitend verbindingen met en via satellieten. De administratieve voorschriften, vnl. Artikel 7 (Regelingen voor de afzonderlijke radiodiensten) en Artikel 9A (Notificeren van frequenties voor stations voor de satelliettelecommunicatie) en de bijbehorende technische criteria, werden op de nieuwe te voorziene situatie afgestemd. Veel aandacht werd besteed aan de maximaal toelaatbare uitgestraalde vermogens, omdat de meeste diensten die van satellietcommunicatie gebruik maken, hun frequenties gezamenlijk, dan wel met aardse diensten zullen moeten delen, waarbij de onderlinge storing tot een aanvaardbaar minimum beperkt moet blijven. Ter voorbereiding van de technische ondergrond voor dit werk vond in februari 1971 een 'Special Joint Meeting' plaats van de betreffende CCIR Studiecommissies.

De frequentietabel van het radioreglement (Artikel 5) onderging een aantal wijzigingen, wat een verruiming van het frequentiegebruik voor satellietcommunicatie mogelijk maakt.

Enige voorbeelden zijn:

1. *De band van 1535 MHz ... 1660 MHz:*
In dit frequentiegebied komen twee banden van 7,5 MHz breedte beschikbaar voor een maritiem mobiel satellietcommunicatiesysteem, twee banden van 15 MHz breedte voor een soortgelijk aëronautisch systeem, en voor gemeenschappelijk gebruik door deze beide diensten twee banden ter breedte van 1 MHz.

2. *Voor omroep via satellieten (voor geluid en beeld):*
– de band van 616 MHz ... 790 MHz voor incidentele toepassingen, nadat door de betrokken landen hiervoor speciale overeenkomsten zullen zijn gesloten;
– de band van 2500 MHz ... 2690 MHz, vnl. voor gebruik buiten Europa;
– de band van 11,7 GHz ... 12,5 GHz. Deze zal met gelijke rechten tevens voor de aardse omroep, het radioverkeer tussen vaste punten en het mobiele radioverkeer beschikbaar zijn, waarbij verder internationaal overleg moet worden afgewacht voor de gedetailleerde bandindeling.

3. *Voor communicatiesatellietsystemen:*
Enige banden ter breedte van 250 MHz tussen 10,7 GHz en 14,5 GHz en twee banden van 3,5 GHz tussen 17,7 GHz en 31 GHz.

Opgemerkt zij nog, dat als nieuwe communicatiedienst werd gedefinieerd die van *satellieten voor exploratie van de aarde*. Hiervoor werd in het frequentiespectrum een plaats aangewezen.

Voor de *radio-astronomie* werd een aantal nieuwe frequentiebandjes vastgelegd. De waarde van het radio-astronomisch onderzoek werd onderkend in de vorm van een aantal voetnoten; de resultaten die behaald kunnen worden blijven echter in belangrijke mate afhankelijk van de mate, waarmee vrijwaring kan worden verkregen van de radiostoring, ondervonden van andere radiocommunicatiediensten.

De wijzigingen in het radioreglement werden door de ca. 100 vertegenwoordigde landen goedgekeurd. Zij zullen per 1 januari 1973 in werking treden.

Ir. F. R. Neubauer.

Elektromagnetische ontstoring en afscherming

Een bijzondere aflevering van 'Siemens Bauteile Informationen', 1970, is gewijd aan ontstoring en afscherming. Men treft er artikelen in aan over eigenschappen van ontstoringsmiddelen en de uitvoering van ontstoringsfilters, over een stoormeetapparaat voor het frequentiegebied van 0,135 ... 3 MHz en over afschermtechnieken.

In één der artikelen, handelend over afschermtechnieken, wordt o.a. vermeld, dat men bij industrieel vervaardigde afgeschermde meetcabines boven 150 kHz een afschermdemping van 80 ... 100 dB kan bereiken. Voor de afscherming wordt veelal koperfolie toegepast, terwijl ventilatieopeningen en vensters van een roosterconstructie zijn voorzien. De doorsnede van de mazen daarin bepaalt boven welke frequentie het rooster doorlatend wordt.

De grootste tot nu toe door Siemens geleverde cabine heeft de afmetingen van 32,03 × 8,09 × 3,38 m. De bovengrens, waarbij deze ruimte nog de gewenste afschermende werking bezit, bedraagt 35 GHz.

Boekennieuws

WALDRON, R. A., **Theory of guided electromagnetic waves**, XVI + 518 blz. Uitg.: Van Nostrand Reinhold Company, London, 1970. Prijs: £ 9.00.

In dit boek worden problemen uit de theorie van de elektromagnetische golfgeleiders behandeld. Dit gebeurt langs theoretische weg en op een niveau dat de schrijver in zijn woord vooraf kenmerkt als liggend in het midden van het spectrum dat van een volledig praktische aan de ene kant tot een mathematisch geavanceerde behandeling aan de andere kant loopt.

Met dit doel voor ogen heeft de schrijver een keuze uit het beschikbare materiaal gedaan, die het beste geïllustreerd wordt door de titels van de hoofdstukken te vermelden: I. The mathematics of waveguide theory, II. Basic electromagnetic theory, III. Transmission lines, IV. Elementary theory of waveguides, V. Cavity resonators, VI. Perturbation theory and its applications, VII. Inhomogeneous and anisotropic waveguides, VIII. Variation theory and its applications, IX. Non-uniform waveguides.

Vooropgesteld moet worden, dat het boek duidelijk is geschreven. De problemen worden helder geformuleerd; dikwijls wordt eerst de oplossingsmethode globaal doorgenomen en wordt daarna een tot in alle details complete afleiding van het resultaat gepresenteerd. Algemene resultaten van zuiver theoretische aard blijven op de achtergrond (zo zal men bijv. reciprociteitsstellingen en de beschrijving van de eigenschappen van een microgolfnetwerk met behulp van de verstrooiingsmatrix niet aantreffen); de aandacht is vooral gericht op het langs theoretische weg analyseren van de eigenschappen van in de microgolftechniek concreet voorkomende configuraties.

Een paar kritische opmerkingen dienen echter te worden geplaatst. Wat benamingen en notaties van de elektromagnetische grootheden betreft zijn er afwijkingen ten aanzien van de internationaal aanvaarde. De elektrische (magnetische) fluxdichtheid $D(B)$ wordt door de schrijver elektrische (magnetische) inductie genoemd; de elektrische ladingsdichtheid wordt aangegeven met q in plaats van ρ en de vector van Poynting met π in plaats van met S. De recensent is verder van mening dat de hoofdstukken I en II zonder enig bezwaar weggelaten hadden kunnen worden, terwijl hoofdstuk III aanmerkelijk korter had kunnen zijn. Iedere lezer die in staat is de rest van het boek te bestuderen, zal met de stof van de eerste drie hoofdstukken ongetwijfeld vertrouwd zijn.

Een belangrijk punt is, dat er in het boek iets misloopt bij de behandeling van de energiestroom in het elektromagnetische veld bij sinusvormig met de tijd veranderende velden. Het is gebruikelijk hierbij de zgn. complexe vector van Poynting in te voeren, waarvan het reële deel in grootte en richting de waarde van de over een periode gemiddelde oppervlaktedichtheid van de vermogensstroom aangeeft. De schrijver weigert (blz. 125) deze complexe vector van Poynting te gebruiken. Dit heeft tot gevolg dat in de energiebeschouwingen nergens duidelijk is of de complexe tijdfactor van de complexe rekenwijze nu wel of niet in de formules is meegenomen en of de vergelijkingen op de reële veldgrootheden, dan wel op hun complexe veldvoorstellingen slaan.

Een belangrijker rol spelen echter in het boek de veldberekeningen. De presentatie hiervan is uitstekend. Er zijn dan ook een aantal resultaten van de hand van de schrijver in het boek te vinden, die anders alleen in de vaktijdschriften te achterhalen zijn.

Samenvattend kan worden gesteld, dat de hoofdstukken die veldberekeningen in golfgeleiderconfiguraties behandelen, een bijdrage tot de vakliteratuur in boekvorm zijn die zeker voor raadpleging in aanmerking komt, zowel door ingenieurs die in dit vakgebied werken, als door studenten die zich kennis van dit vakgebied eigen willen maken.

Prof. dr. ir. A. T. de Hoop.

Uit het NERG

Administratie van het NERG: Postbus 39, Leidschendam. Giro 94746 t.n.v. penningmeester NERG, Leidschendam. Secretariaat van de Examencommissie-NERG: Van Geusaustraat 151, Voorburg.

Ledenmutaties

Nieuwe adressen van leden

Ir. H. G. H. Gooren, Kleine Beerpad 55, Eindhoven.
Dr. ir. F. W. Hogesteeger, Delftweg 147 II, Rijswijk, post Delft.
Ir. H. A. J. M. van Hoof, Peppelschans 24, Leiderdorp.
Ir. A. A. J. Otten, J. A. de Gravenlaan 17, Zoeterwoude.
Ir. W. Werner, Rullen 6, Gerwen (Gem. Nuenen).
Ir. A. J. R. Westbroek, Juliettestraat 17, Amersfoort.